

A Framework for Time-Consistent, Risk-Averse Model Predictive Control: Theory and Algorithms

Yin-Lam Chow¹, Sumeet Singh², Anirudha Majumdar², Marco Pavone²

Abstract

In this paper we present a framework for risk-averse model predictive control (MPC) of linear systems affected by multiplicative uncertainty. Our key innovation is to consider time-consistent, dynamic risk metrics as objective functions to be minimized. This framework is axiomatically justified in terms of time-consistency of risk assessments, is amenable to dynamic optimization, and is unifying in the sense that it captures a full range of risk preferences from risk-neutral to worst case. Within this framework, we propose and analyze an online risk-averse MPC algorithm that is provably stabilizing. Furthermore, by exploiting the dual representation of time-consistent, dynamic risk metrics, we cast the computation of the MPC control law as a convex optimization problem amenable to real-time implementation. Simulation results are presented and discussed.

I. INTRODUCTION

Safety-critical control and decision-making applications demand the consideration of events with small probabilities that can nevertheless have catastrophic effects if realized (e.g., an unmanned aerial vehicle crashing due to an unexpectedly large wind gust or an adaptive cruise control system causing an accident due to a highly unlikely action taken by a neighboring vehicle). Accordingly, one of the main research thrusts for Model Predictive Control (MPC) [1], [2] is to find techniques that are robust in the face of such risks.

Current techniques for handling uncertainty within the MPC framework fall into two categories: (1) min-max (or worst-case) formulations, where the performance indices to be minimized are computed with respect to the worst possible disturbance realization [3], [4], [5], [6], and (2)

¹ Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA. Email: ychow@stanford.edu.

² Department of Aeronautics and Astronautics, Stanford University, CA 94305, USA. Emails: {ssingh19, anirudha, pavone}@stanford.edu.

stochastic formulations, where *risk-neutral expected* values of performance indices (and possibly constraints) are considered [7], [8], [9], [10], [11] (see also the recent review [12]). The main drawback of the worst-case approach is that the control law may be too conservative, since the MPC law is required to guarantee stability and constraint fulfillment under the worst-case scenario (which may have an arbitrarily small probability of occurring). On the other hand, stochastic formulations, whereby the assessment of future random outcomes is accomplished through a risk-neutral expectation, may be unsuitable in scenarios where one desires to protect the system from the risks associated with large deviations.

In general, there are three main challenges with incorporating risk-sensitivity into control and decision-making problems:

Rationality and consistency: The behavior of a control system using a certain risk metric (i.e., a function that maps an uncertain cost to a real number) should be consistent over time. Intuitively, time-consistency stipulates that if a given sequence of costs incurred by the system, when compared to another sequence, has the same current cost and lower risk in the future, then it should be considered less risky at the current time (see Section II-B for a formal statement). Examples of “irrational” behavior that can result from a time-inconsistent risk metric include: (1) a control system intentionally seeking to incur losses [13], or (2) deeming states to be dangerous when in fact they are favorable under any realization of the underlying uncertainty [14], or (3) declaring a decision-making problem to be feasible (e.g., satisfying a certain risk threshold) when in fact it is infeasible under *any* possible subsequent realization of the uncertainties [15]. Remarkably, some of the most common strategies for incorporating risk aversion in decision-making (discussed below) display such inconsistencies [16], [14].

Computational tractability: A risk metric generally adds a nonlinear structure to the optimization problem one must solve in order to compute optimal actions. Hence it is important to ensure the computational tractability of the optimization problem resulting from the choice of risk metric, particularly in dynamic decision-making settings where the control system must plan and react to disturbances in real-time.

Modeling flexibility: One would like to calibrate the risk metric to the control application at hand by: (1) exploring the full spectrum of risk assessments from worst-case to risk-neutral, and (2) ensuring that the risk metric can be applied to a rich set of uncertainty models (e.g., beyond Gaussian models).

Most popular methods in the literature for assessing risks do not satisfy these three require-

ments. The Markowitz mean-variance criterion [17], which has dominated risk management for over 50 years, leads to time-inconsistent assessments of risk in the multi-stage stochastic control framework and also yields computationally intractable problems [13]. Moreover, it is rather limited in terms of modeling flexibility since there is only a single tuning parameter that trades off mean and variance. For example, worst-case risk assessments cannot be captured in such a framework. Finally, the mean-variance metric relies on only the first two moments of the distribution and is thus not well-suited to applications where the disturbance model is non-Gaussian.

A popular alternative to the mean-variance criterion is the entropic risk metric: $\rho(X) = -\log(\mathbb{E}[e^{-\theta X}]) / \theta$, $\theta \in (0, 1)$. The entropic risk metric has been widely studied in the financial mathematics [18], [19] and sequential decision making [20], [21], [22] literatures, and for modeling risk aversion in LQG control problems [23], [24]. While the entropic risk is a more computationally tractable alternative to the mean-variance criterion and can also lead to time-consistent behavior [25], practical applications of the entropic risk metric have proven to be problematic. Notice that the first two terms of the Taylor series expansion of $\rho(X)$ form a weighted sum of mean and variance with regularizer θ , i.e., $\rho(X) \approx \mathbb{E}(X) + \theta \mathbb{E}(X - \mathbb{E}[X])^2$. Consequently, the primary concerns are similar to those associated with the mean-variance measure of risk, e.g., the optimal control policies heavily weight a small number of risk-averse decisions, and are extremely sensitive to errors in the distribution models [26], [27], [28]. The entropic risk metric is a particular example of the general class of methods that model risk aversion using concave utility functions (convex disutility functions in the cost minimization setting). While the expected (dis)utility framework captures the intuitive notion of diminishing marginal utility, it suffers from the issue that even very little risk aversion over moderate costs leads to unrealistically high degrees of risk aversion over large costs [29], [30] (note that this is a limitation of *any* concave utility function). This is an undesirable property from a modeling perspective and thus makes the expected utility framework challenging to apply in practice.

In order to overcome such challenges, in this paper we incorporate risk aversion in MPC by leveraging recent strides in the theory of *dynamic* risk metrics developed by the operations research community [31], [16]. This allows us to propose a framework that satisfies the requirements outlined above with respect to rationality and consistency, computational tractability, and modeling flexibility. Specifically, the key property of dynamic risk metrics is that, by reassessing risk at multiple points in time, one can guarantee time-consistency of risk preferences and the

agent’s behavior [31], [16]. Remarkably, in [16], it is proven that time-consistent risk measures can be represented as a *composition* of one-step risk metrics, which allows for computationally tractable risk evaluation in real-time. Moreover, the one-step risk metrics are *coherent risk metrics* [32], which have been thoroughly investigated and widely applied for static decision-making problems in operations research and finance. Coherent risks were originally conceived in [32] from an *axiomatization* of properties that any rational agent’s risk preferences should satisfy (see Section II for a formal statement of these axioms). In addition to being axiomatically justified, coherent risk metrics capture a wide spectrum of risk assessments from risk neutral to worst-case and thus provide a unifying approach to static risk assessments. Since time-consistent dynamic risks are composed of one-step coherent risks, they inherit the same modeling flexibility.

The contribution of this paper is threefold. First, we introduce a class of dynamic risk metrics, referred to as Markov dynamic polytopic risk metrics, that capture a full range of risk assessments and enjoy a geometrical structure that is particularly favorable from a computational standpoint. Second, we present and analyze a *risk-averse* MPC algorithm that minimizes in a receding-horizon fashion a Markov dynamic polytopic risk metric, under the assumption that the system’s model is linear and is affected by multiplicative uncertainty. Finally, by exploring the *geometric* structure of Markov dynamic polytopic risk metrics, we present a convex programming formulation for risk-averse MPC that is amenable to a real-time implementation (for moderate horizon lengths). Our framework has three main advantages: (1) it is axiomatically justified, in the sense that risk, by construction, is assessed in a time-consistent fashion; (2) it is amenable to dynamic and convex optimization, primarily due to the compositional form of Markov dynamic polytopic risk metrics and their geometry; and (3) it is general, in that it captures a full range of risk assessments from risk-neutral to worst-case. In this respect, our formulation represents a *unifying* approach for risk-averse MPC.

A preliminary version of this paper was presented in [33]. In this extended and revised version, we present the following key extensions: (1) the introduction of constraints on state and control variables for the original infinite-horizon problem, (2) a new offline/online MPC formulation for handling these constraints, (3) a derivation of a computationally verifiable lower bound on the optimal infinite-horizon cost objective and an upper bound on the infinite-horizon cost objective induced by the MPC control policy, (4) additional numerical experimental results including a detailed comparison between our solution algorithm and the one proposed in [8].

The rest of the paper is organized as follows. In Section II we provide a review of the theory

of dynamic risk metrics. In Section III we discuss the stochastic model we address in this paper. In Section IV we introduce and discuss the notion of Markov dynamic polytopic risk metrics. In Section V we state the infinite horizon optimal control problem we wish to address and in Section VI we derive conditions for risk-averse closed-loop stability. In Section VII we present the MPC adaptation of the infinite horizon problem and present various approaches for computation in Section IX. In Section VIII, we derive bounds on the infinite-horizon cost function performance of the optimal and MPC algorithms and thereby rigorously quantify the sub-optimality of our approach. Numerical experiments are presented and discussed in Section X. Finally, in Section XI we draw some conclusions and discuss directions for future work.

II. REVIEW OF DYNAMIC RISK THEORY

In this section, we briefly describe the theory of coherent and dynamic risk metrics, on which we will rely extensively in this paper. The material presented in this section summarizes several novel results in risk theory achieved in the past ten years. Our presentation strives to present this material in an intuitive fashion and with a notation tailored to control applications.

A. Static, Coherent Measures of Risk

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the set of outcomes (sample space), \mathcal{F} is a σ -algebra over Ω representing the set of events we are interested in, and \mathbb{P} is a probability measure over \mathcal{F} . In this paper we will focus on disturbance models characterized by probability *mass* functions, hence we restrict our attention to finite probability spaces (i.e., Ω has a finite number of elements or, equivalently, \mathcal{F} is a finitely generated algebra). Denote with \mathcal{Z} the space of random variables $Z : \Omega \mapsto (-\infty, \infty)$ defined over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In this paper a random variable $Z \in \mathcal{Z}$ is interpreted as a cost, i.e., the smaller the realization of Z , the better. For Z, W , we denote by $Z \leq W$ the point-wise partial order, i.e., $Z(\omega) \leq W(\omega)$ for all $\omega \in \Omega$.

By a *risk measure* (or *risk metric*; we will use these terms interchangeably) we understand a function $\rho(Z)$ that maps an uncertain outcome Z into the extended real line $\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. In this paper we restrict our analysis to *coherent risk measures*, defined as follows:

Definition II.1 (Coherent Risk Measures). *A coherent risk measure is a mapping $\rho : \mathcal{Z} \rightarrow \mathbb{R}$, satisfying the following four axioms:*

- A1 *Convexity:* $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$, for all $\lambda \in [0, 1]$ and $Z, W \in \mathcal{Z}$;
- A2 *Monotonicity:* if $Z \leq W$ and $Z, W \in \mathcal{Z}$, then $\rho(Z) \leq \rho(W)$;
- A3 *Translation invariance:* if $a \in \mathbb{R}$ and $Z \in \mathcal{Z}$, then $\rho(Z + a) = \rho(Z) + a$;
- A4 *Positive homogeneity:* if $\lambda \geq 0$ and $Z \in \mathcal{Z}$, then $\rho(\lambda Z) = \lambda\rho(Z)$.

These axioms were originally conceived in [32] and ensure the “rationality” of single-period risk assessments (we refer the reader to [32] for a detailed motivation of these axioms). One of the main properties for coherent risk metrics is a universal representation theorem for coherent risk metrics, which in the context of *finite* probability spaces takes the following form:

Theorem II.2 (Representation Theorem for Finite Probability Spaces [34, page 265]). *Consider the probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ where Ω is finite and has cardinality $L \in \mathbb{N}$, i.e., $\Omega = \{\omega_1, \dots, \omega_L\}$, \mathcal{F} is the σ -algebra of all subsets (i.e., $\mathcal{F} = 2^\Omega$), and $\mathbb{P} = (p(1), \dots, p(L))$, with all probabilities positive. Let \mathcal{B} be the set of probability density functions: $\mathcal{B} := \left\{ \zeta \in \mathbb{R}^L : \sum_{j=1}^L p(j)\zeta(j) = 1, \zeta \geq 0 \right\}$. The risk measure $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ is a coherent risk measure if and only if there exists a convex bounded and closed set $\mathcal{U} \subset \mathcal{B}$ such that $\rho(Z) = \max_{\zeta \in \mathcal{U}} \mathbb{E}_\zeta[Z]$.*

The result states that any coherent risk measure is an expectation with respect to a worst-case density function ζ , chosen adversarially from a suitable *set* of test density functions (referred to as the *risk envelope*).

B. Dynamic, Time-Consistent Measures of Risk

This section provides a multi-period generalization of the concepts presented in Section II-A and follows closely the discussion in [16]. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a filtration $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \cdots \subset \mathcal{F}_N \subset \mathcal{F}$, and an adapted sequence of real-valued random variables Z_k , $k \in \{0, \dots, N\}$. We assume that $\mathcal{F}_0 = \{\Omega, \emptyset\}$, i.e., Z_0 is deterministic. The variables Z_k can be interpreted as stage-wise costs. For each $k \in \{0, \dots, N\}$, denote with \mathcal{Z}_k the space of random variables defined over the probability space $(\Omega, \mathcal{F}_k, \mathbb{P})$; also, let $\mathcal{Z}_{k,N} := \mathcal{Z}_k \times \cdots \times \mathcal{Z}_N$. Given sequences $Z = \{Z_k, \dots, Z_N\} \in \mathcal{Z}_{k,N}$ and $W = \{W_k, \dots, W_N\} \in \mathcal{Z}_{k,N}$, we interpret $Z \leq W$ component-wise, i.e., $Z_j \leq W_j$ for all $j \in \{k, \dots, N\}$.

The fundamental question in the theory of dynamic risk measures is the following: how do we evaluate the risk of the sequence $\{Z_k, \dots, Z_N\}$ from the perspective of stage k ? The answer, within the modern theory of risk, relies on two key intuitive facts [16]. First, in dynamic settings,

the specification of risk preferences should no longer entail constructing a single risk metric but rather a *sequence* of risk metrics $\{\rho_{k,N}\}_{k=0}^N$, each mapping a future stream of random costs into a risk metric/assessment at time k . This motivates the following definition.

Definition II.3 (Dynamic Risk Measure). *A dynamic risk measure is a sequence of mappings $\rho_{k,N} : \mathcal{Z}_{k,N} \rightarrow \mathcal{Z}_k$, $k \in \{0, \dots, N\}$, obeying the following monotonicity property: $\rho_{k,N}(Z) \leq \rho_{k,N}(W)$ for all $Z, W \in \mathcal{Z}_{k,N}$ such that $Z \leq W$.*

The above monotonicity property (analogous to axiom A2 in Definition II.1) is, arguably, a natural requirement for any meaningful dynamic risk measure.

Second, the sequence of metrics $\{\rho_{k,N}\}_{k=0}^N$ should be constructed so that the risk preference profile is *consistent* over time [35], [36], [37]. A widely accepted notion of time-consistency is as follows [16]: if a certain outcome is considered less risky in all states of the world at stage $k+1$, then it should also be considered less risky at stage k .

The following example (adapted from [15]) shows how dynamic risk measures as defined above might indeed result in *time-inconsistent*, and ultimately undesirable, behaviors.

Example II.4. *Consider the simple setting whereby there is a final cost Z and one seeks to evaluate such a cost from the perspective of earlier stages. Consider the three-stage scenario tree in Figure 1, with the elementary events $\Omega = \{UU, UD, DU, DD\}$, and the filtration $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_1 = \{\emptyset, \{U\}, \{D\}, \Omega\}$, and $\mathcal{F}_2 = 2^\Omega$.*

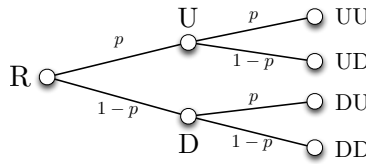


Fig. 1: Scenario tree for example II.4.

Consider the dynamic risk measure:

$$\rho_{k,N}(Z) := \max_{q \in \mathcal{U}} \mathbb{E}_q[Z | \mathcal{F}_k], \quad k = 0, 1, 2$$

where \mathcal{U} contains two probability measures: one corresponding to $p = 0.4$, and the other one to $p = 0.6$. Assume that the random cost is $Z(UU) = Z(DD) = 0$, and $Z(UD) = Z(DU) = 100$. Then, one has $\rho_1(Z)(\omega) = 60$ for all ω , and $\rho_0(Z)(\omega) = 48$. Now consider the following two

options (“policies”). The first option is to receive a (random) cost deduction Z from the three-stage scenario tree above. The second option is to simply receive a deterministic cost deduction $W = 50$ (with no further costs incurred). In this case, the chosen dynamic risk measure deems Z strictly riskier than the deterministic cost W in all states of nature at time $k = 1$, but nonetheless W is deemed riskier than Z at time $k = 0$ – a paradox!

It is important to note that there is nothing special about the selection of this example; similar paradoxical results could be obtained with other risk metrics. We refer the reader to [16], [36], [37] for further insights into the notion of time-consistency and its practical relevance. The issue then is what additional “structural” properties are required for a dynamic risk measure to be time-consistent. We first provide a rigorous version of the previous definition of time-consistency.

Definition II.5 (Time-Consistency ([16])). *A dynamic risk measure $\{\rho_{k,N}\}_{k=0}^N$ is time-consistent if, for all $0 \leq l < k \leq N$ and all sequences $Z, W \in \mathcal{Z}_{l,N}$, the conditions*

$$Z_i = W_i, \quad i = l, \dots, k-1, \quad \text{and} \quad \rho_{k,N}(Z_k, \dots, Z_N) \leq \rho_{k,N}(W_k, \dots, W_N),$$

imply that

$$\rho_{l,N}(Z_l, \dots, Z_N) \leq \rho_{l,N}(W_l, \dots, W_N).$$

As we will see in Theorem II.7, the notion of time-consistent risk measures is tightly linked to the notion of coherent risk measures, whose generalization to the multi-period setting is given below:

Definition II.6 (Coherent One-step Conditional Risk Measures ([16])). *A coherent one-step conditional risk measure is a mapping $\rho_k : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k$, $k \in \{0, \dots, N-1\}$, with the following four properties:*

- *Convexity:* $\rho_k(\lambda Z + (1 - \lambda)W) \leq \lambda \rho_k(Z) + (1 - \lambda) \rho_k(W)$, $\forall \lambda \in [0, 1]$ and $Z, W \in \mathcal{Z}_{k+1}$;
- *Monotonicity:* if $Z \leq W$ then $\rho_k(Z) \leq \rho_k(W)$, $\forall Z, W \in \mathcal{Z}_{k+1}$;
- *Translation invariance:* $\rho_k(Z + W) = Z + \rho_k(W)$, $\forall Z \in \mathcal{Z}_k$ and $W \in \mathcal{Z}_{k+1}$;
- *Positive homogeneity:* $\rho_k(\lambda Z) = \lambda \rho_k(Z)$, $\forall Z \in \mathcal{Z}_{k+1}$ and $\lambda \geq 0$.

We are now in a position to state the main result of this section.

Theorem II.7 (Dynamic, Time-consistent Risk Measures ([16])). *Consider, for each $k \in \{0, \dots, N\}$, the mappings $\rho_{k,N} : \mathcal{Z}_{k,N} \rightarrow \mathcal{Z}_k$ defined as*

$$\rho_{k,N} = Z_k + \rho_k(Z_{k+1} + \rho_{k+1}(Z_{k+2} + \dots + \rho_{N-2}(Z_{N-1} + \rho_{N-1}(Z_N)) \dots)), \quad (1)$$

where the ρ_k 's are coherent one-step conditional risk measures. Then, the ensemble of such mappings is a dynamic, time-consistent risk measure.

Remarkably, Theorem 1 in [16] shows (under weak assumptions) that the “multi-stage composition” in equation (1) is indeed *necessary for time-consistency*. Accordingly, in the remainder of this paper, we will focus on the *dynamic, time-consistent risk measures* characterized in Theorem II.7.

III. MODEL DESCRIPTION

Consider the discrete time system:

$$x_{k+1} = A(w_k)x_k + B(w_k)u_k, \quad (2)$$

where $k \in \mathbb{N}$ is the time index, $x_k \in \mathbb{R}^{N_x}$ is the state, $u_k \in \mathbb{R}^{N_u}$ is the (unconstrained) control input, and $w_k \in \mathcal{W}$ is the process disturbance. We assume that the initial condition x_0 is deterministic. We assume that \mathcal{W} is a finite set of cardinality L , i.e., $\mathcal{W} = \{w^{[1]}, \dots, w^{[L]}\}$. For each stage k and state-control pair (x_k, u_k) , the process disturbance w_k is drawn from set \mathcal{W} according to the probability mass function $p = [p(1), p(2), \dots, p(L)]^\top$, where $p(j) = \mathbb{P}(w_k = w^{[j]})$, $j \in \{1, \dots, L\}$. Without loss of generality, we assume that $p(j) > 0$ for all j . Note that the probability mass function for the process disturbance is time-invariant, and that the process disturbance is *independent* of the process history and of the state-control pair (x_k, u_k) . Under these assumptions, the stochastic process $\{x_k\}$ is clearly a Markov process.

By enumerating all realizations of the process disturbance w_k , system (2) can be rewritten as:

$$x_{k+1} = \begin{cases} A_1 x_k + B_1 u_k & \text{if } w_k = w^{[1]}, \\ \vdots & \vdots \\ A_L x_k + B_L u_k & \text{if } w_k = w^{[L]}, \end{cases}$$

where $A_j := A(w^{[j]})$ and $B_j := B(w^{[j]})$, $j \in \{1, \dots, L\}$.

The results presented in this paper can be immediately extended to the time-varying case (i.e., where the probability mass function for the process disturbance is time-varying). To simplify notation, however, we prefer to focus this paper on the time-invariant case.

IV. MARKOV POLYTOPIC RISK MEASURES

In this section we *refine* the notion of dynamic time-consistent risk metrics (as defined in Theorem II.7) in two ways: (1) we add a polytopic structure to the dual representation of coherent risk metrics, and (2) we add a Markovian structure. This will lead to the definition of Markov dynamic polytopic risk metrics, which enjoy favorable computational properties and, at the same time, maintain most of the generality of dynamic time-consistent risk metrics.

A. Polytopic Risk Measures

According to the discussion in Section III, the probability space for the process disturbance has a finite number of elements. Accordingly, consider Theorem II.2; by definition of expectation, one has $\mathbb{E}_\zeta[Z] = \sum_{j=1}^L Z(j)p(j)\zeta(j)$. In our framework (inspired by [38]), we consider coherent risk measures where the risk envelope \mathcal{U} is a *polytope*, i.e., there exist matrices S^I, S^E and vectors T^I, T^E of appropriate dimensions such that

$$\mathcal{U}^{\text{poly}} = \{\zeta \in \mathcal{B} \mid S^I \zeta \leq T^I, S^E \zeta = T^E\}.$$

We will refer to coherent risk measures representable with a polytopic risk envelope as *polytopic risk measures*. Consider the bijective map $q(j) := p(j)\zeta(j)$ (recall that, in our model, $p(j) > 0$). Then, by applying such map, one can easily rewrite a polytopic risk measure as

$$\rho(Z) = \max_{q \in \mathcal{U}^{\text{poly}}} \mathbb{E}_q[Z],$$

where q is a *probability mass function* belonging to a polytopic subset of the standard simplex, i.e.:

$$\mathcal{U}^{\text{poly}} = \left\{ q \in \Delta^L \mid S^I q \leq T^I, S^E q = T^E \right\}, \quad (3)$$

where $\Delta^L := \{q \in \mathbb{R}^L : \sum_{j=1}^L q(j) = 1, q \geq 0\}$. Accordingly, one has $E_q[Z] = \sum_{j=1}^L Z(j)q(j)$ (note that, with a slight abuse of notation, we are using the same symbols as before for $\mathcal{U}^{\text{poly}}$, S^I , and S^E).

The class of polytopic risk measures is large: we give below some examples (also note that any comonotonic risk measure is a polytopic risk measure [37]).

Example IV.1 (Examples of Polytopic Risk Measures). *As a first example, the expected value of a random variable Z can be represented according to equation (3) with polytopic risk envelope*

$$\mathcal{U}^{\text{poly}} = \left\{ q \in \Delta^L \mid q(j) = p(j), j \in \{1, \dots, L\} \right\}.$$

A second example is represented by the average upper semi-deviation risk metric, defined as

$$\rho_{\text{AUS}}(Z) := \mathbb{E}[Z] + c \mathbb{E}[(Z - \mathbb{E}[Z])^+],$$

where $0 \leq c \leq 1$ and $(x)^+ := \max(0, x)$. This metric can be represented according to equation (3) with polytopic risk envelope ([39], [34]):

$$\mathcal{U}^{\text{poly}} = \left\{ q \in \Delta^L \mid q(j) = p(j) \left(1 + h(j) - \sum_{j=1}^L h(j)p(j) \right), 0 \leq h(j) \leq c, j \in \{1, \dots, L\} \right\}.$$

A related risk metric is the mean absolute semi-deviation risk metric, defined as

$$\rho_{\text{AS}}(Z) = \mathbb{E}[Z] + c \mathbb{E} \left[\left| Z - \mathbb{E}[Z] \right| \right],$$

where $0 \leq c \leq 1$. This metric can be represented using a risk envelope identical to that for ρ_{AUS} with the only change being $h(j) \in [-c, c]$ [39].

A risk metric that is very popular in the finance industry is the Conditional Value-at-Risk (CVaR), defined as ([40])

$$\text{CVaR}_\alpha(Z) := \inf_{y \in \mathbb{R}} \left[y + \frac{1}{\alpha} \mathbb{E}[(Z - y)^+] \right], \quad (4)$$

where $\alpha \in (0, 1]$. CVaR_α can be represented according to equation (3) with the polytopic risk envelope (see [34]):

$$\mathcal{U}^{\text{poly}} = \left\{ q \in \Delta^L \mid 0 \leq q(j) \leq \frac{p(j)}{\alpha}, j \in \{1, \dots, L\} \right\}.$$

As a special case, CVaR_0 corresponds to the worst case risk and can be trivially represented according to (3) with polytopic risk envelope $\mathcal{U}^{\text{poly}} = \Delta^L$.

Other important examples include spectral risk measures [41], optimized certainty equivalent and expected utility [42], [34], and distributionally-robust risk [8]. The key point is that the notion of polytopic risk metric covers a full gamut of risk assessments, ranging from risk-neutral to worst case.

B. Markov Dynamic Polytopic Risk Metrics

Note that in the definition of dynamic, time-consistent risk measures, since at stage k the value of ρ_k is \mathcal{F}_k -measurable, the evaluation of risk can depend on the *whole* past, see [16, Section IV]. For example, the weight c in the definition of the average upper mean semi-deviation risk metric can be an \mathcal{F}_k -measurable random variable (see [16, Example 2]). This generality, which appears

of little practical value in many cases, leads to optimization problems that are intractable. This motivates us to add a *Markovian structure* to dynamic, time-consistent risk measures (similarly as in [16]). We start by introducing the notion of Markov polytopic risk measure (similar to [16, Definition 6]).

Definition IV.2 (Markov Polytopic Risk Measures). *Consider the Markov process $\{x_k\}$ that evolves according to equation (2). A coherent one-step conditional risk measure $\rho_k(\cdot)$ is a Markov polytopic risk measure with respect to $\{x_k\}$ if it can be written as*

$$\rho_k(Z(x_{k+1})) = \max_{q \in \mathcal{U}_k^{\text{poly}}(x_k, p)} \mathbb{E}_q[Z(x_{k+1})]$$

where $\mathcal{U}_k^{\text{poly}}(x_k, p) = \{q \in \Delta^L \mid S_k^I(x_k, p)q \leq T_k^I(x_k, p), S_k^E(x_k, p)q = T_k^E(x_k, p)\}$ is the polytopic risk envelope.

In other words, a Markov polytopic risk measure is a coherent one-step conditional risk measure where the evaluation of risk is not allowed to depend on the whole past (for example, the weight c in the definition of the average upper mean semi-deviation risk metric can depend on the past only through x_k), and the risk envelope is a polytope. Correspondingly, we define a Markov dynamic polytopic risk metric as follows.

Definition IV.3 (Markov Dynamic Polytopic Risk Measures). *Consider the Markov process $\{x_k\}$ that evolves according to equation (2). A Markov dynamic polytopic risk measure is a set of mappings $\rho_{k,N} : \mathcal{Z}_{k,N} \rightarrow \mathcal{Z}_k$ defined as*

$$\rho_{k,N} = Z(x_k) + \rho_k(Z(x_{k+1}) + \dots + \rho_{N-2}(Z(x_{N-1}) + \rho_{N-1}(Z(x_N))) \dots),$$

for $k \in \{0, \dots, N\}$, where ρ_k are single-period Markov polytopic risk measures.

Clearly, a Markov dynamic polytopic risk metric is time-consistent. Definition IV.3 can be extended to the case where the probability distribution for the disturbance depends on the current state and control action. We avoid this generalization to keep the exposition simple and consistent with model (2).

V. PROBLEM FORMULATION

In light of Sections III and IV, we are now in a position to state the risk-averse optimization problem we wish to solve in this paper. Our problem formulation relies on Markov dynamic polytopic risk metrics that satisfy the following stationarity assumption.

Assumption V.1 (Time-invariance of Risk Assessments). *The polytopic risk envelopes $\mathcal{U}_k^{\text{poly}}$ are independent of time k and state x_k , i.e. $\mathcal{U}_k^{\text{poly}}(x_k, p) = \mathcal{U}^{\text{poly}}(p)$, $\forall k$.*

This assumption is crucial for the well-posedness of our formulation and to devise a tractable MPC algorithm that relies on linear matrix inequalities. We next introduce a notion of stability tailored to our risk-averse context.

Definition V.2 (Uniform Global Risk-Sensitive Exponential Stability). *System (2) is said to be Uniformly Globally Risk-Sensitive Exponentially Stable (UGRSES) if there exist constants $c \geq 0$ and $\lambda \in [0, 1)$ such that for all initial conditions $x_0 \in \mathbb{R}^{N_x}$,*

$$\rho_{0,k}(0, \dots, 0, x_k^\top x_k) \leq c \lambda^k x_0^\top x_0, \quad \text{for all } k \in \mathbb{N}, \quad (5)$$

where $\{\rho_{0,k}\}$ is a Markov dynamic polytopic risk measure satisfying Assumption V.1. If condition (5) only holds for initial conditions within some bounded neighborhood Ω of the origin, the system is said to be Uniformly Locally Risk-Sensitive Exponentially Stable (ULRSES) with domain Ω .

Note that, in general, UGRSES is a *more restrictive* stability condition than mean-square stability, as illustrated by the following example.

Example V.3 (Mean-Square Stability versus Risk-Sensitive Stability). *System (2) is said to be Uniformly Globally Mean-Square Exponentially Stable (UGMSES) if there exist constants $c \geq 0$ and $\lambda \in [0, 1)$ such that for all initial conditions $x_0 \in \mathbb{R}^{N_x}$,*

$$\mathbb{E} [x_k^\top x_k] \leq c \lambda^k x_0^\top x_0, \quad \text{for all } k \in \mathbb{N},$$

see [43, Definition 1] and [8, Definition 1]. Consider the discrete time system

$$x_{k+1} = \begin{cases} \sqrt{0.5} x_k & \text{with probability } 0.2, \\ \sqrt{1.1} x_k & \text{with probability } 0.8. \end{cases} \quad (6)$$

A sufficient condition for system (6) to be UGMSES is that there exist positive definite matrices $P = P^\top \succ 0$ and $L = L^\top \succ 0$ such that

$$\mathbb{E} [x_{k+1}^\top P x_{k+1}] - x_k^\top P x_k \leq -x_k^\top L x_k,$$

for all $k \in \mathbb{N}$, see [8, Lemma 1]. One can easily check that with $P = 100$ and $L = 1$ the above inequality is satisfied, and, hence system (6) is UGMSES.

Assuming risk is assessed according to the Markov dynamic polytopic risk metric $\rho_{0,k} = \text{CVaR}_{0.5} \circ \dots \circ \text{CVaR}_{0.5}$, we next show that system (6) is not UGRSES. In fact, using the dual representation given in Example IV.1, one can write

$$\text{CVaR}_{0.5}(Z(x_{k+1})) = \max_{q \in \mathcal{U}^{\text{poly}}} \mathbb{E}_q[Z(x_{k+1})],$$

where $\mathcal{U}^{\text{poly}} = \{q \in \Delta^2 \mid 0 \leq q_1 \leq 0.4, 0 \leq q_2 \leq 1.6\}$. Consider the probability mass function $\bar{q} = [0.1/1.1, 1/1.1]^\top$. Since $\bar{q} \in \mathcal{U}^{\text{poly}}$, one has

$$\text{CVaR}_{0.5}(x_{k+1}^2) \geq 0.5 x_k^2 \frac{0.1}{1.1} + 1.1 x_k^2 \frac{1}{1.1} = 1.0455 x_k^2.$$

By repeating this argument, one can then show that

$$\rho_{0,k}(x_{k+1}^2) = \text{CVaR}_{0.5} \circ \dots \circ \text{CVaR}_{0.5}(x_{k+1}^2) \geq a^{k+1} x_0^\top x_0,$$

where $a = 1.0455$. Hence, one cannot find constants c and λ that satisfy equation (5). Consequently, system (6) is UGMSES but not UGRSES.

Consider the MDP described in Section III and let Π be the set of all stationary feedback control policies, i.e., $\Pi := \{\pi : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_u}\}$. Consider the quadratic cost function $C : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \rightarrow \mathbb{R}_{\geq 0}$ defined as $C(x, u) := \|x\|_Q^2 + \|u\|_R^2$, where $Q = Q^\top \succ 0$ and $R = R^\top \succ 0$ are given state and control penalties, and $\|x\|_A^2$ defines the weighted norm, i.e., $x^\top A x$. Define the multi-stage cost function:

$$J_{0,k}(x_0, \pi) := \rho_{0,k}(C(x_0, \pi(x_0)), \dots, C(x_k, \pi(x_k))),$$

where $\{\rho_{0,k}\}$ is a Markov dynamic polytopic risk measure satisfying Assumption V.1. The problem we wish to address is as follows.

Optimization Problem \mathcal{OPT} — Given an initial state $x_0 \in \mathbb{R}^{N_x}$, solve

$$\begin{aligned} & \inf_{\pi \in \Pi} \quad \limsup_{k \rightarrow \infty} J_{0,k}(x_0, \pi) \\ & \text{subject to} \quad x_{k+1} = A(w_k)x_k + B(w_k)\pi(x_k) \\ & \quad \|T_u \pi(x_k)\|_2 \leq u_{\max}, \quad \|T_x x_k\|_2 \leq x_{\max} \\ & \quad \text{System is UGRSES} \end{aligned}$$

where (T_u, u_{\max}) and (T_x, x_{\max}) describe constraints on the control and state respectively (given as second-order cone constraints).

We denote the optimal cost function as $J_{0,\infty}^*(x_0)$. Note that the risk measure in the definition of UGRSES is assumed to be identical to the risk measure used to evaluate the cost of a policy. Also, by Assumption V.1, the single-period risk metrics are time-invariant, hence one can write

$$\begin{aligned} \rho_{0,k}\left(C(x_0, \pi(x_0)), \dots, C(x_k, \pi(x_k))\right) &= C(x_0, \pi(x_0)) \\ &\quad + \rho(C(x_1, \pi(x_1)) + \dots + \rho(C(x_k, \pi(x_k))) \dots), \end{aligned} \quad (7)$$

where ρ is a given Markov polytopic risk metric that models the “amount” of risk aversion. This paper addresses problem \mathcal{OPT} along three main dimensions:

- 1) Find sufficient conditions for *risk-sensitive* stability (i.e., for UGRSES).
- 2) Design a MPC algorithm to efficiently compute a suboptimal state-feedback control policy.
- 3) Find lower bounds for the optimal cost of problem \mathcal{OPT} .

VI. RISK-SENSITIVE STABILITY

In this section we provide a sufficient condition for system (2) to be UGRSES, under the assumptions of Section V. This condition relies on Lyapunov techniques and is inspired by [8] (Lemma VI.1 indeed reduces to Lemma 1 in [8] when the risk measure is simply an expectation).

Lemma VI.1 (Sufficient Conditions for UGRSES). *Consider a policy $\pi \in \Pi$ and the corresponding closed-loop dynamics for system (2), denoted by $x_{k+1} = f(x_k, w_k)$. The closed-loop system is UGRSES if there exists a function $V(x) : \mathbb{R}^{N_x} \rightarrow \mathbb{R}$ and scalars $b_1, b_2, b_3 > 0$, such that for all $x \in \mathbb{R}^{N_x}$,*

$$\begin{aligned} b_1 \|x\|^2 &\leq V(x) \leq b_2 \|x\|^2, \text{ and} \\ \rho(V(f(x, w))) - V(x) &\leq -b_3 \|x\|^2. \end{aligned} \quad (8)$$

We call such a function $V(x)$ a risk-sensitive Lyapunov function.

Proof. From the time-consistency, monotonicity, translational invariance, and positive homogeneity of Markov dynamic polytopic risk measures, condition (8) implies

$$\begin{aligned} \rho_{0,k+1}(0, \dots, 0, b_1 \|x_{k+1}\|^2) &\leq \rho_{0,k+1}(0, \dots, 0, V(x_{k+1})) \\ &= \rho_{0,k}(0, \dots, 0, V(x_k) + \rho(V(x_{k+1}) - V(x_k))) \\ &\leq \rho_{0,k}(0, \dots, 0, V(x_k) - b_3 \|x_k\|^2) \\ &\leq \rho_{0,k}(0, \dots, 0, (b_2 - b_3) \|x_k\|^2). \end{aligned}$$

Also, since $\rho_{0,k+1}$ is monotonic, one has $b_1\rho_{0,k+1}(0, \dots, 0, \|x_{k+1}\|^2) \geq 0$, which implies $b_2 \geq b_3$ and in turn $(1 - b_3/b_2) \in [0, 1)$. Since $V(x_k)/b_2 \leq \|x_k\|^2$, by using the previous inequalities one can write:

$$\rho_{0,k+1}(0, \dots, 0, V(x_{k+1})) \leq \rho_{0,k}(0, \dots, 0, V(x_k) - b_3\|x_k\|^2) \leq \left(1 - \frac{b_3}{b_2}\right) \rho_{0,k}(0, \dots, 0, V(x_k)).$$

Repeating this bounding process, one obtains:

$$\begin{aligned} \rho_{0,k+1}(0, \dots, 0, V(x_{k+1})) &\leq \left(1 - \frac{b_3}{b_2}\right)^k \rho_{0,1}(V(x_1)) = \left(1 - \frac{b_3}{b_2}\right)^k \rho(V(x_1)) \\ &\leq \left(1 - \frac{b_3}{b_2}\right)^k (V(x_0) - b_3\|x_0\|^2) \leq b_2 \left(1 - \frac{b_3}{b_2}\right)^{k+1} \|x_0\|^2. \end{aligned}$$

Again, by monotonicity, the above result implies

$$\rho_{0,k+1}(0, \dots, 0, x_{k+1}^\top x_{k+1}) \leq \frac{b_2}{b_1} \left(1 - \frac{b_3}{b_2}\right)^{k+1} x_0^\top x_0.$$

By setting $c = b_2/b_1$ and $\lambda = (1 - b_3/b_2) \in [0, 1)$, the claim is proven. \square

Remark VI.2 (Sufficient Conditions for ULRSES). *The closed-loop system is ULRSES with domain Ω if the conditions in (8) only hold within the bounded set Ω .*

VII. MODEL PREDICTIVE CONTROL PROBLEM

This section describes a MPC strategy that approximates the solution to \mathcal{OPT} . We note that while an exact solution to \mathcal{OPT} would lead to time-consistent risk assessments, MPC is not guaranteed to be time consistent over an entire infinite horizon realization since it is inherently myopic. The MPC strategy thus provides an efficiently implementable policy that approximately mimics the time-consistent nature of the optimal solution to \mathcal{OPT} .

A. The Unconstrained Case

In this section we set up the receding horizon version of problem \mathcal{OPT} under the assumption that there are no constraints. Consider the following receding-horizon cost function for $N \geq 1$:

$$\begin{aligned} J(x_{k|k}, \pi_{k|k}, \dots, \pi_{k+N-1|k}, P) &:= \rho_{k,k+N}(C(x_{k|k}, \pi_{k|k}(x_{k|k})), \dots, \\ &\quad C(x_{k+N-1|k}, \pi_{k+N-1|k}(x_{k+N-1|k}), x_{k+N}^T P x_{k+N})), \end{aligned} \tag{9}$$

where $x_{h|k}$ is the state at time h predicted at stage k (a *discrete* random variable), $\pi_{h|k}$ is the control *policy* to be applied at time h as determined at stage k (i.e., $\pi_{h|k} : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_u}$), and

$P = P^T \succ 0$ is a terminal weight matrix. We are now in a position to state the model predictive control problem.

Optimization problem \mathcal{MPC} — Given an initial state $x_{k|k} \in \mathbb{R}^{N_x}$ and a prediction horizon $N \geq 1$, solve

$$\begin{aligned} \min_{\pi_{k|k}, \dots, \pi_{k+N-1|k}} \quad & J(x_{k|k}, \pi_{k|k}, \dots, \pi_{k+N-1|k}, P) \\ \text{subject to} \quad & x_{k+h+1|k} = A(w_{k+h})x_{k+h|k} + B(w_{k+h})\pi_{k+h|k}(x_{k+h|k}), h \in \{0, \dots, N-1\}. \end{aligned}$$

Note that a Markov policy is guaranteed to be optimal for problem \mathcal{MPC} (see [16, Theorem 2]). The optimal cost function for problem \mathcal{MPC} is denoted by $J_k^*(x_{k|k})$, and a minimizing policy is denoted by $\{\pi_{k|k}^*, \dots, \pi_{k+N-1|k}^*\}$ (if multiple minimizing policies exist, then one of the minimizing policies is selected arbitrarily). For each state x_k , we set $x_{k|k} = x_k$ and the (time-invariant) model predictive control law is then defined as

$$\pi^{MPC}(x_k) = \pi_{k|k}^*(x_{k|k}). \quad (10)$$

Note that the model predictive control problem \mathcal{MPC} involves an optimization over *time-varying closed-loop policies*, as opposed to the classical deterministic case where the optimization is over open-loop sequences. A similar approach is taken in [7], [8]. We will show in Section IX how to solve problem \mathcal{MPC} efficiently.

The following theorem shows that the model predictive control law (10), with a proper choice of the terminal weight P , is risk-sensitive stabilizing, i.e., the closed-loop system (2) is UGRSES.

Theorem VII.1 (Stochastic Stability for Model Predictive Control Law, Unconstrained Case). *Consider the model predictive control law in equation (10) and the corresponding closed-loop dynamics for system (2) with initial condition $x_0 \in \mathbb{R}^{N_x}$. Suppose that $P = P^T \succ 0$, and there exists a matrix F such that:*

$$\sum_{j=1}^L q_l(j) (A_j + B_j F)^T P (A_j + B_j F) - P + Q + F^T R F \prec 0, \quad (11)$$

for all $l \in \{1, \dots, \text{card}(\mathcal{U}^{\text{poly}, V}(p))\}$, where $\mathcal{U}^{\text{poly}, V}(p)$ is the set of vertices of polytope $\mathcal{U}^{\text{poly}}(p)$, q_l is the l th element in set $\mathcal{U}^{\text{poly}, V}(p)$, and $q_l(j)$ denotes the j th component of vector q_l , $j \in \{1 \dots, L\}$. Then, the closed loop system (2) is UGRSES.

Proof. We will show that J_k^* is a risk-sensitive Lyapunov function (Lemma VI.1). Specifically, we want to show that J_k^* satisfies the two inequalities in equation (8); the claim then follows by simply noting that, in our time-invariant setup, J_k^* does not depend on k .

Consider the bottom inequality in equation (8). At time k consider problem \mathcal{MPC} with state $x_{k|k}$. The sequence of optimal control policies is given by $\{\pi_{k+h|k}^*\}_{h=0}^{N-1}$. Let us define a sequence of control policies from time $k+1$ to N according to

$$\pi_{k+h|k+1}(x_{k+h|k}) := \begin{cases} \pi_{k+h|k}^*(x_{k+h|k}) & \text{if } h \in [1, N-1], \\ F x_{k+N|k} & \text{if } h = N. \end{cases} \quad (12)$$

This is simply the concatenation of the sequence $\{\pi_{k+h|k}^*\}_{h=1}^{N-1}$ with a linear feedback law for stage N (the reason why we refer to this policy with the subscript “ $k+h|k+1$ ” is that we will use this policy as a feasible policy for problem \mathcal{MPC} starting at stage $k+1$).

Consider problem \mathcal{MPC} at stage $k+1$ with initial condition given by $x_{k+1|k+1} = A(w_k)x_{k|k} + B(w_k)\pi_{k|k}^*(x_{k|k})$, and denote with $\bar{J}_{k+1}(x_{k+1|k+1})$ the cost of the objective function for the \mathcal{MPC} problem corresponding to the control policy sequence in (12). Note that $x_{k+1|k+1}$ (and therefore $\bar{J}_{k+1}(x_{k+1|k+1})$) is a random variable with L possible realizations, given $x_{k|k}$. Define:

$$Z_{k+N} := x_{k+N|k}^T (-P + Q + F^T R F) x_{k+N|k},$$

$$Z_{k+N+1} := ((A(w_{k+N|k}) + B(w_{k+N|k})F)x_{k+N|k})^T P ((A(w_{k+N|k}) + B(w_{k+N|k})F)x_{k+N|k}).$$

By exploiting the dual representation of Markov polytopic risk metrics, one can write

$$\begin{aligned} Z_{k+N} + \rho_{k+N}(Z_{k+N+1}) &= x_{k+N|k}^T (-P + Q + F^T R F) x_{k+N|k} \\ &+ \max_{q \in \mathcal{U}^{\text{poly}}(p)} \sum_{j=1}^L q(j) x_{k+N|k}^T (A_j + B_j F)^T P (A_j + B_j F) x_{k+N|k}. \end{aligned}$$

Combining the equation above with equation (11), one readily obtains the inequality

$$Z_{k+N} + \rho_{k+N}(Z_{k+N+1}) \leq 0. \quad (13)$$

One can then write the following chain of inequalities:

$$\begin{aligned}
J_k^*(x_{k|k}) &= C(x_{k|k}, \pi_{k|k}^*(x_{k|k})) + \rho_k \left(\rho_{k+1,N} \left(C(x_{k+1|k}, \pi_{k+1|k}^*(x_{k+1|k})), \dots, \right. \right. \\
&\quad \left. \left. \|x_{k+N|k}\|_Q^2 + \|x_{k+N|k}\|_{F^T R F}^2 + \rho_{k+N}(Z_{k+N+1}) - Z_{k+N} - \rho_{k+N}(Z_{k+N+1}) \right) \right) \\
&\geq C(x_{k|k}, \pi_{k|k}^*(x_{k|k})) + \rho_k \left(\rho_{k+1,N} \left(C(x_{k+1|k}, \pi_{k+1|k}^*(x_{k+1|k})), \dots, \right. \right. \\
&\quad \left. \left. \|x_{k+N|k}\|_Q^2 + \|x_{k+N|k}\|_{F^T R F}^2 + \rho_{k+N}(Z_{k+N+1}) \right) \right) \\
&= C(x_{k|k}, \pi_{k|k}^*(x_{k|k})) + \rho_k \left(\bar{J}_{k+1}(x_{k+1|k+1}) \right) \\
&\geq C(x_{k|k}, \pi_{k|k}^*(x_{k|k})) + \rho_k \left(J_{k+1}^*(x_{k+1|k+1}) \right), \tag{14}
\end{aligned}$$

where the first equality follows from the definitions of Z_{k+N} and of dynamic, time-consistent risk measures, the second inequality follows from equation (13) and the monotonicity property of Markov polytopic risk metrics (see also [16, Page 242]), the third equality follows from the definition of $\bar{J}_{k+1}(x_{k+1|k+1})$, and the fourth inequality follows from the definition of J_{k+1}^* and the monotonicity of Markov polytopic risk metrics.

Consider now the top inequality in equation (8). One can easily bound $J_k^*(x_{k|k})$ from below according to:

$$J_k^*(x_{k|k}) \geq x_{k|k}^T Q x_{k|k} \geq \lambda_{\min}(Q) \|x_{k|k}\|^2, \tag{16}$$

where $\lambda_{\min}(Q) > 0$ by assumption. To bound $J_k^*(x_{k|k})$ from above, define:

$$M_A := \max_{r \in \{0, \dots, N-1\}} \max_{j_0, \dots, j_r \in \{1, \dots, L\}} \|A_{j_r} \dots A_{j_1} A_{j_0}\|_2.$$

Since the problem is unconstrained (and, hence, zero is a feasible control input) and by exploiting the monotonicity property, one can write:

$$\begin{aligned}
J_k^*(x_{k|k}) &\leq C(x_{k|k}, 0) + \rho_k \left(C(x_{k+1|k}, 0) + \rho_{k+1} \left(C(x_{k+2|k}, 0) + \dots + \rho_{k+N-1} (\|x_{k+N|k}\|_P^2) \dots \right) \right) \\
&\leq \|Q\|_2 \|x_{k|k}\|_2^2 + \rho_k \left(\|Q\|_2 \|x_{k+1|k}\|_2^2 + \rho_{k+1} \left(\|Q\|_2 \|x_{k+2|k}\|_2^2 + \dots \right. \right. \\
&\quad \left. \left. + \rho_{k+N-1} (\|P\|_2 \|x_{k+N|k}\|_2^2) \dots \right) \right).
\end{aligned}$$

Therefore, by using the translational invariance and monotonicity property of Markov polytopic risk measures, one obtains the upper bound:

$$J_k^*(x_{k|k}) \leq (N \|Q\|_2 + \|P\|_2) M_A \|x_{k|k}\|_2^2. \tag{17}$$

Combining the results in equations (15), (16), (17), and given the time-invariance of our problem setup, one concludes that $J_k^*(x_{k|k})$ is a risk-sensitive Lyapunov function for the closed-loop system (2), in the sense of Lemma VI.1. This concludes the proof. \square

B. The Constrained Case

We now enforce the state and control constraints introduced in problem \mathcal{OPT} within the receding horizon framework. Consider the time-invariant ellipsoids:

$$\mathbb{U} := \{u \in \mathbb{R}^{N_u} \mid \|T_u u\|_2 \leq u_{\max}\}, \quad \mathbb{X} := \{x \in \mathbb{R}^{N_x} \mid \|T_x x\|_2 \leq x_{\max}\}.$$

While we focus on ellipsoidal state and control constraints in this paper, our methodology can readily accommodate component-wise and polytopic constraints via suitable LMI representations, for example, see [44], [45] for detailed derivations.

Our receding horizon framework may be decomposed into two steps. First, offline, we search for an ellipsoidal set \mathcal{E}_{\max} and a *local* feedback control law $u(x) = Fx$ that renders \mathcal{E}_{\max} control invariant and ensures satisfaction of state and control constraints. Additionally, within the offline step, we search for a terminal cost matrix P (for the online MPC problem) to ensure that the closed-loop dynamics under the model predictive controller are risk-sensitive exponentially stable. The online MPC optimization then constitutes the second step of our framework. Consider first, the offline step. We parameterize \mathcal{E}_{\max} as follows:

$$\mathcal{E}_{\max}(W) := \{x \in \mathbb{R}^{N_x} \mid x^\top W^{-1} x \leq 1\}, \quad (18)$$

where W (and hence W^{-1}) is a positive definite matrix. The (offline) optimization problem to solve for W , F , and P is presented below.

Optimization Problem \mathcal{PE} — Solve

$$\begin{aligned} & \max_{W=W^\top \succ 0, F, P=P^\top \succ 0} \log \det(W) \\ & \text{subject to} \quad F^\top \frac{T_u^\top T_u}{u_{\max}^2} F - W^{-1} \preceq 0, \end{aligned} \quad (19)$$

$$\begin{aligned} & \sum_{j=1}^L q_l(j) (A_j + B_j F)^\top P (A_j + B_j F) - P \\ & + (F^\top R F + Q) \prec 0, \forall q_l \in \mathcal{U}^{\text{poly}, V}(p) \end{aligned} \quad (20)$$

$$\forall j \in \{1, \dots, L\} :$$

$$(A_j + B_j F)^\top \frac{T_x^\top T_x}{x_{\max}^2} (A_j + B_j F) - W^{-1} \preceq 0, \quad (21)$$

$$(A_j + B_j F)^\top W^{-1} (A_j + B_j F) - W^{-1} \preceq 0. \quad (22)$$

The objective in problem \mathcal{PE} is to maximize the volume of the control invariant ellipsoid $\mathcal{E}_{\max}(W)$. Note that $\mathcal{E}_{\max}(W)$ may contain states outside of \mathbb{X} , however, we restrict our domain of interest to the intersection $\mathbb{X} \cap \mathcal{E}_{\max}(W)$. The semi-definite inequality in (20) defines the terminal cost matrix P (ref. inequality (11) in Theorem VII.1), and will be instrumental in proving risk-sensitive stability for system (2) under the model predictive control law. Note that inequality (20) is *bi-linear* in the decision variables. In Section IX, we will derive an equivalent Linear Matrix Inequality (LMI) characterization of (20) in order to derive efficient solution algorithms. We first analyze the properties of the state feedback control law $u(x) = Fx$ within the set $\mathcal{E}_{\max}(W)$.

Lemma VII.2 (Properties of \mathcal{E}_{\max}). *Suppose problem \mathcal{PE} is feasible and $x \in \mathbb{X} \cap \mathcal{E}_{\max}(W)$. Let $u(x) = Fx$. Then, the following statements are true:*

- 1) $\|T_u u\|_2 \leq u_{\max}$, i.e., the control constraint is satisfied.
- 2) $\|T_x (A(w)x + B(w)u)\|_2 \leq x_{\max}$ surely, i.e., the state constraint is satisfied at the next step surely.
- 3) $A(w)x + B(w)u \in \mathcal{E}_{\max}(W)$ surely, i.e., the set $\mathcal{E}_{\max}(W)$ is robust control invariant under the control law $u(x)$.

Thus, $u(x) \in \mathbb{U}$ and $A(w)x + B(w)u \in \mathbb{X} \cap \mathcal{E}_{\max}(W)$ surely.

Proof. We first prove 1) and 2) and thereby establish $u(x)$ as a feasible control law within the set $\mathcal{E}_{\max}(W)$. Notice that:

$$\|T_u Fx\|_2 \leq u_{\max} \Leftrightarrow \|T_u F W^{\frac{1}{2}} (W^{-\frac{1}{2}} x)\|_2 \leq u_{\max}.$$

From (18), applying the Schur complement, we know that $\|W^{-\frac{1}{2}}x\|_2 \leq 1$ for any $x \in \mathcal{E}_{\max}(W)$. Thus, by the Cauchy Schwarz inequality, a *sufficient* condition is given by $\|T_u F W^{\frac{1}{2}}\|_2 \leq u_{\max}$, which can be written as

$$(F W^{\frac{1}{2}})^{\top} T_u^{\top} T_u (F W^{\frac{1}{2}}) \preceq u_{\max}^2 I \Leftrightarrow F^{\top} T_u^{\top} T_u F \preceq u_{\max}^2 W^{-1}.$$

Re-arranging the inequality above yields the expression given in (19). The state constraint can be proved in an identical fashion by leveraging (18) and (21). It is omitted for brevity.

We now prove the third statement. By definition of a robust control invariant set, we are required to show that for any $x \in \mathcal{E}_{\max}(W)$, that is, for all x that satisfy the inequality: $x^{\top} W^{-1} x \leq 1$, application of the control law $u(x)$ yields the following inequality:

$$(A_j x + B_j F x)^{\top} W^{-1} (A_j x + B_j F x) \leq 1, \forall j \in \{1, \dots, L\}.$$

Using the S-procedure [46], it is equivalent to show that there exists $\lambda \geq 0$ such that the following condition holds:

$$\begin{bmatrix} \lambda W^{-1} - (A_j + B_j F)^{\top} W^{-1} (A_j + B_j F) & 0 \\ * & 1 - \lambda \end{bmatrix} \succeq 0,$$

for all $j \in \{1, \dots, L\}$. By setting $\lambda = 1$, one obtains the largest feasibility set for W and F . The expression in (22) corresponds to the (1,1) block in the matrix above. \square

Lemma VII.2 establishes $\mathbb{X} \cap \mathcal{E}_{\max}(W)$ as a robust control invariant set under the feasible feedback control law $u(x) = Fx$. This result will be crucial to ascertain the persistent feasibility and closed-loop stability properties of the online optimization algorithm.

We are now ready to formalize the MPC problem. Suppose the feasible set of solutions in problem \mathcal{PE} is non-empty and define $W = W^*$ and $P = P^*$, where W^*, P^* are the maximizers for problem \mathcal{PE} . Consider the following online optimization problem:

Optimization problem \mathcal{MPC} — Given an initial state $x_{k|k} \in \mathbb{X}$ and a prediction horizon $N \geq 1$, solve

$$\begin{aligned} \min_{\pi_{k|k}, \dots, \pi_{k+N-1|k}} \quad & J(x_{k|k}, \pi_{k|k}, \dots, \pi_{k+N-1|k}, P) \\ \text{subject to} \quad & x_{k+h+1|k} = A(w_{k+h})x_{k+h|k} + B(w_{k+h})\pi_{k+h|k}(x_{k+h|k}), \\ & \pi_{k+h|k}(x_{k+h|k}) \in \mathbb{U}, \quad x_{k+h+1|k} \in \mathbb{X}, \quad h \in \{0, \dots, N-1\}, \\ & x_{k+N|k} \in \mathcal{E}_{\max}(W) \text{ surely.} \end{aligned}$$

Note that a Markov policy is guaranteed to be optimal for problem \mathcal{MPC} (see [16, Theorem 2]). The optimal cost function for problem \mathcal{MPC} is denoted by $J_k^*(x_{k|k})$, and a minimizing policy is denoted by $\{\pi_{k|k}^*, \dots, \pi_{k+N-1|k}^*\}$. For each state x_k , we set $x_{k|k} = x_k$ and the (time-invariant) model predictive control law is then defined as

$$\pi^{MPC}(x_k) = \pi_{k|k}^*(x_{k|k}). \quad (23)$$

Note that problem \mathcal{MPC} involves an optimization over *time-varying closed-loop policies*, as opposed to the classical deterministic case where the optimization is over open-loop control inputs. A similar approach is taken in [7], [8]. We will show in Section IX how to solve problem \mathcal{MPC} efficiently. We now address the persistent feasibility and stability properties for problem \mathcal{MPC} .

Theorem VII.3 (Persistent Feasibility). *Define \mathcal{X}_N to be the set of initial states for which problem \mathcal{MPC} is feasible. Assume $x_{k|k} \in \mathcal{X}_N$ and the control law is given by (23). Then, it follows that $x_{k+1|k+1} \in \mathcal{X}_N$ surely.*

Proof. Given $x_{k|k} \in \mathcal{X}_N$, problem \mathcal{MPC} may be solved to yield a closed-loop optimal control policy:

$$\{\pi_{k|k}^*(x_{k|k}), \dots, \pi_{k+N-1|k}^*(x_{k+N-1|k})\},$$

such that $x_{k+N|k} \in \mathbb{X} \cap \mathcal{E}_{\max}(W)$. Consider problem \mathcal{MPC} at stage $k+1$ with initial condition $x_{k+1|k+1}$. From Lemma VII.2, we know that the set $\mathbb{X} \cap \mathcal{E}_{\max}(W)$ is robust control invariant under the feasible feedback control law $u(x) = Fx$. Thus,

$$\{\pi_{k+1|k}^*(x_{k+1|k}), \dots, \pi_{k+N-1|k}^*(x_{k+N-1|k}), Fx_{k+N|k}\}, \quad (24)$$

is a feasible control policy at stage $k+1$. Note that this is simply a concatenation of the optimal tail policy from the previous iteration $\{\pi_{k+h|k}^*(x_{k+h|k})\}_{h=1}^{N-1}$, with the state feedback law $Fx_{k+N|k}$ for the final step.

Since a feasible control policy exists at stage $k+1$, $x_{k+1|k+1} = A_j x_{k|k} + B_j \pi_{k|k}^*(x_{k|k}) \in \mathcal{X}_N$ for any $j \in \{1, \dots, L\}$, completing the proof. \square

Theorem VII.4 (Stochastic Stability with MPC). *Suppose the initial state x_0 lies within \mathcal{X}_N . Then, under the model predictive control law given in (23), the closed-loop system is ULRSES with domain \mathcal{X}_N .*

Proof. The first part of the proof is identical to the reasoning presented in the proof for Theorem VII.1. In particular, we leverage the policy given in (24) as a feasible policy for problem \mathcal{MPC} at stage $k + 1$ and inequality (20) to show:

$$J_k^*(x_{k|k}) \geq C(x_{k|k}, \pi_{k|k}^*(x_{k|k})) + \rho_k(J_{k+1}^*(x_{k+1|k+1})), \quad (25)$$

for all $x_{k|k} \in \mathcal{X}_N$. Additionally, we retain the same lower bound for $J_k^*(x_{k|k})$ as given in (16). The upper bound for $J_k^*(x_{k|k})$ is derived in two steps. First, define

$$M_A := \max_{r \in \{0, \dots, N-1\}} \max_{j_0, \dots, j_r \in \{1, \dots, L\}} \alpha_{j_r} \dots \alpha_{j_1} \alpha_{j_0},$$

where $\alpha_j := \|A_j + B_j F\|_2$.

Suppose $x_{k|k} \in \mathbb{X} \cap \mathcal{E}_{\max}(W)$. From Lemma VII.2, we know that the control policy $\pi_{k+h|k}(x_{k+h|k}) = \{F x_{k+h|k}\}_{h=0}^{N-1}$ is feasible and consequently, $\mathbb{X} \cap \mathcal{E}_{\max}(W) \subseteq \mathcal{X}_N$. Defining $\theta_f := \|Q + F^\top R F\|_2$, we thus have

$$\begin{aligned} J_k^*(x_{k|k}) &\leq C(x_{k|k}, F x_{k|k}) + \rho_k \left(C(x_{k+1|k}, F x_{k+1|k}) + \dots + \rho_{k+N-1} (x_{k+N|k}^\top P x_{k+N|k}) \dots \right) \\ &\leq \theta_f \|x_{k|k}\|_2^2 + \rho_k \left(\theta_f \|x_{k+1|k}\|_2^2 + \dots + \rho_{k+N-1} (\|P\|_2 \|x_{k+N|k}\|_2^2) \dots \right), \end{aligned}$$

for all $x_{k|k} \in \mathbb{X} \cap \mathcal{E}_{\max}(W)$. Exploiting the translational invariance and monotonicity property of Markov polytopic risk metrics, one obtains the upper bound:

$$J_k^*(x_{k|k}) \leq \underbrace{(N \theta_f + \|P\|_2)}_{:= \beta > 0} M_A \|x_{k|k}\|_2^2, \quad \forall x_{k|k} \in \mathbb{X} \cap \mathcal{E}_{\max}(W). \quad (26)$$

In order to derive an upper bound for $J_k^*(x_{k|k})$ with the above structure for all $x_{k|k} \in \mathcal{X}_N$, we draw inspiration from a similar proof in [47]. By leveraging the finite cardinality of the disturbance set \mathcal{W} and the set closure preservation property attributed to the inverse of continuous functions, it is possible to show that \mathcal{X}_N is closed. Then, since \mathcal{X}_N is necessarily a subset of the bounded set \mathbb{X} , it follows that \mathcal{X}_N is compact. Thus, there exists some constant $\Gamma > 0$ such that $J_k^*(x_{k|k}) \leq \Gamma$ for all $x_{k|k} \in \mathcal{X}_N$. That Γ is finite follows from the fact that $\{\|x_{k+h|k}\|_2\}_{h=0}^N$ and $\{\|\pi_{k+h|k}(x_{k+h|k})\|_2\}_{h=0}^{N-1}$ are finitely bounded for all $x_{k|k} \in \mathcal{X}_N$. Now since $\mathcal{E}_{\max}(W)$ is compact and non-empty, there exists a $d > 0$ such that $\mathcal{E}_d := \{x \in \mathbb{R}^{N_x} \mid \|x\|_2 \leq d\} \subset \mathcal{E}_{\max}(W)$. Let $\hat{\beta} = \max\{\beta \|x\|_2^2 \mid \|x\|_2 \leq d\}$. Consider now, the function: $(\Gamma/\hat{\beta})\beta \|x\|_2^2$. Then since $\beta \|x\|_2^2 > \hat{\beta}$ for all $x \in \mathcal{X}_N \setminus \mathcal{E}_d$ and $\Gamma \geq \hat{\beta}$, it follows that

$$J_k^*(x_{k|k}) \leq \left(\frac{\Gamma \beta}{\hat{\beta}} \right) \|x_{k|k}\|_2^2, \quad \forall x_{k|k} \in \mathcal{X}_N, \quad (27)$$

as desired. Combining the results in equations (25), (16), (27), and given the time-invariance of our problem setup, one concludes that $J_k^*(x_{k|k})$ is a risk-sensitive Lyapunov function for the closed-loop system (2), in the sense of Lemma VI.1. This concludes the proof. \square

Remark VII.5 (Performance Comparisons). *The two-step optimization methodology proposed via Problems \mathcal{PE} and \mathcal{MPC} is similar to the approach described in [8] in that both the control invariant ellipsoid (\mathcal{E}_{\max}) and the conditions to ensure stability are computed offline, while problem \mathcal{MPC} is solved online. This hybrid procedure is more computationally efficient than the online algorithm given in [48], and boasts better performance as compared with the offline algorithm in [3]. On the other hand, the stability analysis here differs from [8] since we use J_k^* as the Lyapunov function instead of the fixed quadratic form described in [8]. This allows us to formulate a less-constrained framework and achieve superior performance. To gain additional insight into this comparison, we present an alternative formulation of problems \mathcal{PE} and \mathcal{MPC} in Appendix A, analogous to the approach in [8], and present numerical results exemplifying the performance improvement in Section X-B.*

VIII. BOUNDS ON OPTIMAL COST

In this section, by leveraging semi-definite programming, we provide a lower bound for the optimal cost of problem \mathcal{OPT} and an upper bound for the optimal cost using the MPC algorithm. These bounds will be used in Section X to bound the factor of sub-optimality for our MPC control algorithm. In the following, let

$$\bar{A} := \begin{bmatrix} A_1^\top & \dots & A_L^\top \end{bmatrix}^\top, \quad \text{and} \quad \bar{B} := \begin{bmatrix} B_1^\top & \dots & B_L^\top \end{bmatrix}^\top,$$

and for each $q_l \in \mathcal{U}^{\text{poly},V}(p)$, define

$$\Sigma_l := \text{diag}(q_l(1), \dots, q_l(L)) \succ 0.$$

Theorem VIII.1 (Bounds for Problem \mathcal{OPT}). *Lower Bound: Suppose the feasible set of the following Linear Matrix Inequality in the symmetric matrix $X \succ 0$ is non-empty:*

$$\begin{bmatrix} R + \bar{B}^\top (\Sigma_l \otimes X) \bar{B} & \bar{B}^\top (\Sigma_l \otimes X) \bar{A} \\ * & \bar{A}^\top (\Sigma_l \otimes X) \bar{A} - (X - Q) \end{bmatrix} \succeq 0, \quad (28)$$

for all $l \in \{1, \dots, \text{card}(\mathcal{U}^{\text{poly},V}(p))\}$. Then, the optimal cost of problem \mathcal{OPT} can be lower bounded as

$$J_{0,\infty}^*(x_0) \geq \max\{x_0^\top X x_0 : X \text{ satisfies LMI (28)}\}.$$

Upper Bound: For all $x_0 \in \mathcal{X}_N$,

$$J_0^*(x_0) \geq \limsup_{k \rightarrow \infty} J_{0,k}(x_0, \pi^{MPC}).$$

Proof. Lower Bound: Consider a symmetric matrix X such that the LMI in equation (28) is satisfied. Also, let π be a stationary feedback control policy that is feasible for problem \mathcal{OPT} . At stage k , consider a state x_k (reachable under policy π) and the corresponding control action $u_k = \pi(x_k)$ (since π is a feasible policy, the pair (x_k, u_k) clearly satisfies the state-control constraints). By pre- and post-multiplying the LMI (28) with $[u_k^\top, x_k^\top]$ and its transpose, one obtains the inequality

$$u_k^\top R u_k + u_k^\top \bar{B}^\top (\Sigma_l \otimes X) \bar{B} u_k + 2u_k^\top \bar{B}^\top (\Sigma_l \otimes X) \bar{A} x_k + x_k^\top (\bar{A}^\top (\Sigma_l \otimes X) \bar{A} - (X - Q)) x_k \geq 0,$$

which implies

$$x_k^\top X x_k - \sum_{j=1}^L q_l(j) (A_j x_k + B_j u_k)^\top X (A_j x_k + B_j u_k) \leq u_k^\top R u_k + x_k^\top Q x_k,$$

for all $l \in \{1, \dots, \text{cardinality}(\mathcal{U}^{\text{poly}, V}(p))\}$. Since $\mathcal{U}^{\text{poly}}(p)$ is a convex polytope of probability vectors q (with vertex set $\mathcal{U}^{\text{poly}, V}(p)$), then the inequality above holds for any $q \in \mathcal{U}^{\text{poly}}(p)$. Exploiting the dual representation of Markov polytopic risk measures, one has $\rho_k(x_{k+1}^\top X x_{k+1}) = \max_{q \in \mathcal{U}^{\text{poly}}(p)} E_q[x_{k+1}^\top X x_{k+1}]$, which leads to the inequality

$$x_k^\top X x_k - \rho_k(x_{k+1}^\top X x_{k+1}) \leq u_k^\top R u_k + x_k^\top Q x_k.$$

As the above inequality holds for all $k \in \mathbb{N}$, one can write, for all $k \in \mathbb{N}$,

$$\sum_{h=0}^k (x_h^\top X x_h - \rho_h(x_{h+1}^\top X x_{h+1})) \leq \sum_{h=0}^k (u_h^\top R u_h + x_h^\top Q x_h).$$

Since each single-period risk measure is monotone, their composition $\rho_0 \circ \dots \circ \rho_{k-1}$ is monotone as well. Hence by applying $\rho_0 \circ \dots \circ \rho_{k-1}$ to both sides one obtains

$$\rho_0 \circ \dots \circ \rho_{k-1} \left(\sum_{h=0}^k (x_h^\top X x_h - \rho_h(x_{h+1}^\top X x_{h+1})) \right) \leq \rho_0 \circ \dots \circ \rho_{k-1} \left(\sum_{h=0}^k (u_h^\top R u_h + x_h^\top Q x_h) \right).$$

By repeatedly applying the translational invariance property (see Definition II.6), the right-hand side can be written as

$$\begin{aligned} & \|u_0\|_R^2 + \|x_0\|_Q^2 + \rho_0(\|u_1\|_R^2 + \|x_1\|_Q^2 + \dots + \rho_{k-1}(\|u_k\|_R^2 + \|x_k\|_Q^2) \dots) \\ &= \rho_{0,k} (u_0^\top R u_0 + x_0^\top Q x_0, \dots, u_k^\top R u_k + x_k^\top Q x_k) = J_{0,k}(x_0, \pi), \end{aligned}$$

where the last equality follows from the definition of dynamic, time-consistent risk measures (Theorem II.7). As for the left-hand side, note that the translation invariance and positive homogeneity property imply that a coherent one-step conditional risk measure is subadditive, i.e., $\rho_h(Z + W) \leq \rho_h(Z) + \rho_h(W)$, where $Z, W \in \mathcal{Z}_{h+1}$. In turn, subadditivity implies that $\rho_h(Z - W) \geq \rho_h(Z) - \rho_h(W)$. Hence, by repeatedly applying the translation invariance and monotonicity property and the inequality $\rho_h(Z - W) \geq \rho_h(Z) - \rho_h(W)$, one obtains

$$\begin{aligned}
& \rho_0 \circ \dots \circ \rho_{k-1} \left(\sum_{h=0}^k (\|x_h\|_X^2 - \rho_h(\|x_{h+1}\|_X^2)) \right) \\
&= \|x_0\|_X^2 - \rho_0(\|x_1\|_X^2) + \rho_0 \left(\|x_1\|_X^2 - \rho_1(\|x_2\|_X^2) + \rho_1(\|x_2\|_X^2 - \rho_2(\|x_3\|_X^2) + \dots \right. \\
&\quad \left. + \rho_{k-1}(\|x_k\|_X^2 - \rho_k(\|x_{k+1}\|_X^2) \dots) \right) \\
&\geq \|x_0\|_X^2 - \rho_0 \circ \dots \circ \rho_{k-1} \circ \rho_k(x_{k+1}^\top X x_{k+1}) \\
&= x_0^\top X x_0 - \rho_{0,k+1}(0, \dots, x_{k+1}^\top X x_{k+1}).
\end{aligned}$$

Since π is a feasible policy, then $\lim_{k \rightarrow \infty} \rho_{0,k}(0, \dots, x_{k+1}^\top X x_{k+1}) = 0$ almost surely. Hence, one readily obtains (using the monotonicity and positive homogeneity property)

$$\lim_{k \rightarrow \infty} \rho_{0,k+1}(0, \dots, 0, x_{k+1}^\top X x_{k+1}) \leq \lambda_{\max}(X) \lim_{k \rightarrow \infty} \rho_{0,k+1}(0, \dots, x_{k+1}^\top X x_{k+1}) = 0,$$

almost surely. Collecting all results so far, one has the inequality

$$x_0^\top X x_0 \leq \lim_{k \rightarrow \infty} J_{0,k}(x_0, \pi),$$

for all symmetric matrices satisfying the LMI (28) and all feasible policies π . By maximizing the left-hand side and minimizing the right-hand side one obtains the claim.

Upper Bound: For all $k \in \mathbb{N}$, inequality (15) provides the relation

$$J_k^*(x_{k|k}) \geq C(x_{k|k}, \pi^{MPC}(x_{k|k})) + \rho_k(J_{k+1}^*(x_{k+1|k+1})).$$

Since $x_0 \in \mathcal{X}_N$ and problem \mathcal{MPC} is recursively feasible, we obtain the following sequence of state-control pairs: $\{(x_{k|k}, \pi^{MPC}(x_{k|k}))\}_{k=0}^\infty$. Applying inequality (15) recursively and using the

monotonicity property of coherent one-step risk measures, we deduce the following:

$$\begin{aligned}
J_0^*(x_{0|0}) &\geq C(x_{0|0}, \pi^{MPC}(x_{0|0})) + \rho_0(J_1^*(x_{1|1})) \\
&\geq C(x_{0|0}, \pi^{MPC}(x_{0|0})) + \rho_0(C(x_{1|1}, \pi^{MPC}(x_{1|1})) + \rho_1(J_2^*(x_{2|2}))) \\
&\geq \dots \geq C(x_{0|0}, \pi^{MPC}(x_{0|0})) + \rho_0\left(C(x_{1|1}, \pi^{MPC}(x_{1|1})) + \dots \right. \\
&\quad \left. + \rho_{k-1}(C(x_{k|k}, \pi^{MPC}(x_{k|k})) + J_{k+1}^*(x_{k+1|k+1})) \dots \right) \\
&\geq C(x_{0|0}, \pi^{MPC}(x_{0|0})) + \rho_0(C(x_{1|1}, \pi^{MPC}(x_{1|1})) + \dots + \rho_{k-1}(C(x_{k|k}, \pi^{MPC}(x_{k|k}))) \dots) \\
&= \rho_{0,k}(C(x_{0|0}, \pi^{MPC}(x_{0|0})), \dots, C(x_{k|k}, \pi^{MPC}(x_{k|k}))), \forall k,
\end{aligned}$$

where the second to last inequality follows from the fact that $J_{k+1}^*(x_{k+1|k+1}) \geq 0$, and the equality follows from the definition of dynamic, time-consistent risk metrics. Noting that $x_{k|k} = x_k$ for all $k \in \mathbb{N}$ and by taking the limit $k \rightarrow \infty$ on both sides, one obtains the claim. \square

IX. SOLUTION ALGORITHMS

Prior to solving problem \mathcal{MPC} , one would first need to find a matrix P that satisfies (20). Expression (20) is a bilinear semi-definite inequality in (P, F) . It is well known that checking feasibility of a bilinear semi-definite inequality constraint is an NP-hard problem [49]. Nonetheless, one can transform this bilinear semi-definite inequality constraint into an LMI by applying the Projection Lemma [50]. The next two results present LMI characterizations of conditions (19), (20), (21), and (22). The proofs are provided in Appendix B.

Theorem IX.1 (LMI Characterization of Stability Constraint). *Consider the following set of LMIs with decision variables Y , G , $\overline{Q} = \overline{Q}^\top \succ 0$:*

$$\begin{bmatrix} I_{L \times L} \otimes \overline{Q} & 0 & 0 & -\Sigma_l^{\frac{1}{2}}(\overline{A}G + \overline{B}Y) \\ * & R^{-1} & 0 & -Y \\ * & * & I & -Q^{\frac{1}{2}}G \\ * & * & * & -\overline{Q} + G + G^\top \end{bmatrix} \succ 0, \quad (29)$$

for all $l \in \{1, \dots, \text{card}(\mathcal{U}^{\text{poly}, V}(p))\}$. The expression in (20) is equivalent to the set of LMIs in (29) by setting $F = YG^{-1}$ and $P = \overline{Q}^{-1}$.

Furthermore, by the application of the Projection Lemma to the expressions in (19), (21) and (22), we obtain the following corollary:

Corollary IX.2. Suppose the following set of LMIs with decision variables Y , G , and $W = W^\top \succ 0$ are satisfied:

$$\begin{bmatrix} x_{\max}^2 I & -T_x(A_j G + B_j Y) \\ * & -W + G + G^\top \end{bmatrix} \succ 0, \begin{bmatrix} u_{\max}^2 I & -T_u Y \\ * & -W + G + G^\top \end{bmatrix} \succ 0, \begin{bmatrix} W & -(A_j G + B_j Y) \\ * & -W + G + G^\top \end{bmatrix} \succ 0. \quad (30)$$

Then, by setting $F = YG^{-1}$, the LMIs above represent sufficient conditions for the LMIs in (19), (21) and (22).

Note that in Corollary IX.2, strict inequalities are imposed only for the sake of analytical simplicity when applying the Projection Lemma. Using similar arguments as in [3], non-strict versions of the above LMIs may also be derived, for example, leveraging some additional technicalities [51].

A solution approach for the receding horizon adaptation of problem \mathcal{OPT} is to first solve the LMIs in Theorem IX.1 and Corollary IX.2. If a solution for (P, Y, G, W) is found, problem \mathcal{MPC} can be solved via dynamic programming (see [16, Theorem 2]) after state and action *discretization*, see, e.g., [52], [53]. Note that the discretization process might yield a large-scale dynamic programming problem for which the computational complexity scales exponentially with the resolution of discretization. This motivates the convex programming approach presented next.

A. Convex Programming Approach

While problem \mathcal{MPC} is defined as an optimization over *Markov* control policies, in the convex programming approach, we re-define the problem as an optimization over *history-dependent* policies. One can show (with a virtually identical proof) that the stability Theorem VII.4 still holds when history-dependent policies are considered. Furthermore, since Markov policies are optimal in our setup (see [16, Theorem 2]), the value of the optimal cost stays the same. The key advantage of history-dependent policies is that their additional flexibility leads to a convex formulation of the online problem. Consider the following parameterization of *history-dependent* control policies. Let $j_0, \dots, j_h \in \{1, \dots, L\}$ be the realized indices for the disturbances in the first $h+1$ steps of the \mathcal{MPC} problem, where $h \in \{1, \dots, N-1\}$. The control to be exerted at stage h is denoted by $\bar{U}_h(j_0, \dots, j_{h-1})$. Similarly, we refer to the state at stage h as $\bar{X}_h(j_0, \dots, j_{h-1})$.

The dependence on (j_0, \dots, j_{h-1}) enables us to keep track of the growth of the scenario tree.

In terms of this new notation, the system dynamics (2) can be rewritten as:

$$\begin{aligned}\bar{X}_0 &:= x_{k|k}, \quad \bar{U}_0 \in \mathbb{U}, \quad \bar{X}_1(j_0) = A_{j_0} \bar{X}_0 + B_{j_0} \bar{U}_0, \quad h = 1, \\ \bar{X}_h(j_0, \dots, j_{h-1}) &= A_{j_{h-1}} \bar{X}_{h-1}(j_0, \dots, j_{h-2}) + B_{j_{h-1}} \bar{U}_{h-1}(j_0, \dots, j_{h-2}), \quad h \geq 2.\end{aligned}\quad (31)$$

The final solution algorithm is presented below.

Algorithm MPC — Given an initial state $x_0 \in \mathbb{X}$ and a prediction horizon $N \geq 1$, solve

- **Offline step:** Solve

$$\max_{W=W^\top \succ 0, G, Y, \bar{Q}=\bar{Q}^\top \succ 0} \log \det(W)$$

subjected to the LMIs in expressions (29) and (30).

- **Online Step:** Suppose the feasible set of solutions in the offline step is non-empty. Define: $W = W^*$ and $P = (\bar{Q}^*)^{-1}$ where W^* and \bar{Q}^* are the maximizers for the offline step. Now at each step $k \in \{0, 1, \dots\}$, solve:

$$\begin{aligned}\min_{\gamma_2(j_0, \dots, j_{N-1}), \bar{X}_h(j_0, \dots, j_{h-1}), \bar{U}_0,} \quad & \rho_{k,k+N}(C(x_{k|k}, \bar{U}_0), \dots, C(\bar{X}_{N-1}, \bar{U}_{N-1}), \gamma_2) \\ & \bar{U}_h(j_0, \dots, j_{h-1}), h \in \{1, \dots, N\}, \\ & j_0, \dots, j_{N-1} \in \{1, \dots, L\}\end{aligned}\quad (32)$$

subject to

- the LMIs

$$\begin{bmatrix} 1 & \bar{X}_N(j_0, \dots, j_{N-1})^\top \\ * & W \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_2(j_0, \dots, j_{N-1})I & \bar{X}_N(j_0, \dots, j_{N-1})^\top \\ * & P^{-1} \end{bmatrix} \succeq 0. \quad (33)$$

- the system dynamics in equation (31);
- the control and state constraints for $h \in \{1, \dots, N\}$:

$$\begin{aligned}\|T_u \bar{U}_0\|_2 &\leq u_{\max}, \quad \|T_u \bar{U}_h(j_0, \dots, j_{h-1})\|_2 \leq u_{\max}, \\ \|T_x \bar{X}_h(j_0, \dots, j_{h-1})\|_2 &\leq x_{\max}\end{aligned}\quad (34)$$

Then, set $\pi^{MPC}(x_{k|k}) = \bar{U}_0$.

Note that the terminal cost has been *equivalently* reformulated via the epigraph constraint in

(33) using the variable γ_2 in order to represent it within the LMI framework (in contrast to the original representation of the cost, which was quadratic in the decision variables). This algorithm is clearly suitable only for “moderate” values of N , given the combinatorial explosion of the scenario tree. As a degenerate case, when we exclude all lookahead steps, problem \mathcal{MPC} is reduced to an *offline* optimization. By trading off performance, one can compute the control policy offline and implement it directly online without further optimization:

Algorithm \mathcal{MPC}^0 — Given $x_0 \in \mathbb{X}$, solve:

$$\min_{\gamma_2, W = W^\top \succ 0, G, Y, \bar{Q} = \bar{Q}^\top \succ 0} \gamma_2$$

subject to LMIs (29), (30) and

$$\begin{bmatrix} 1 & x_0^\top \\ * & W \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \gamma_2 I & x_0^\top \\ * & \bar{Q} \end{bmatrix} \succeq 0.$$

Then, set $\pi^{\mathcal{MPC}}(x_k) = YG^{-1}x_k$.

The domain of feasibility for \mathcal{MPC}^0 is the control invariant set $\mathbb{X} \cap \mathcal{E}_{\max}(W)$. Showing ULRSSES for algorithm \mathcal{MPC}^0 is more straightforward than the corresponding analysis for problem \mathcal{MPC} and is summarized within the following corollary.

Corollary IX.3 (Quadratic Lyapunov Function). *Suppose problem \mathcal{MPC}^0 is feasible. Then, system (2) under the offline MPC policy: $\pi^{\mathcal{MPC}}(x_k) = YG^{-1}x_k$ is ULRSSES with domain $\mathbb{X} \cap \mathcal{E}_{\max}(W)$.*

Proof. From Theorem IX.1, we know that the set of LMIs in (29) is equivalent to the expression in (20) when $F = YG^{-1}$. Then since $x_0 \in \mathbb{X} \cap \mathcal{E}_{\max}(W)$, a robust control invariant set under the local feedback control law $u(x) = YG^{-1}x$, exploiting the dual representation of Markov polytopic risk measures yields the inequality

$$\rho_k(x_{k+1}^\top P x_{k+1}) - x_k^\top P x_k \leq -x_k^\top L x_k \quad \forall k \in \mathbb{N}, \quad (35)$$

where $L = Q + (YG^{-1})^\top R (YG^{-1}) = L^\top \succ 0$. Define the Lyapunov function $V(x) = x^\top P x$. Set $b_1 = \lambda_{\min}(P) > 0$, $b_2 = \lambda_{\max}(P) > 0$ and $b_3 = \lambda_{\min}(L) > 0$. Then by Lemma VI.1, this stochastic system is ULRSSES with domain $\mathbb{X} \cap \mathcal{E}_{\max}(W)$. \square

Note that our algorithms require a vertex representation of the risk polytopes (rather than the hyperplane representation in Definition IV.2). In our implementation, we use the vertex enumeration function included in the MPT toolbox [54], which relies on the simplex method.

X. NUMERICAL EXPERIMENTS

In this section we present several numerical experiments that were run on a 2.3 GHz Intel Core i5, MacBook Pro laptop, using the MATLAB YALMIP Toolbox (version 2.6.3 [55]) with the SDPT3 solver. All measurements of computation time are given in seconds.

A. A 2-state, 2-input Stochastic System

Consider a stochastic system with 6 scenarios: $x_{k+1} = A(w_k)x_k + B(w_k)u_k$ where $w_k \in \{1, 2, 3, 4, 5, 6\}$ and

$$\begin{aligned} A_1 &= \begin{bmatrix} 2 & 0.5 \\ -0.5 & 2 \end{bmatrix}, A_2 = \begin{bmatrix} -0.1564 & -0.0504 \\ -0.0504 & -0.1904 \end{bmatrix}, A_3 = \begin{bmatrix} 1.5 & -0.3 \\ 0.2 & 1.5 \end{bmatrix}, A_4 = \begin{bmatrix} 0.5768 & 0.2859 \\ 0.2859 & 0.7499 \end{bmatrix}, \\ A_5 &= \begin{bmatrix} 1.8 & 0.3 \\ -0.2 & 1.8 \end{bmatrix}, A_6 = \begin{bmatrix} 0.2434 & 0.3611 \\ 0.3611 & 0.3630 \end{bmatrix}, \\ B_1 &= \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, B_2 = \begin{bmatrix} -0.9540 & 0 \\ -0.7773 & 0.1852 \end{bmatrix}, B_3 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, B_4 = \begin{bmatrix} -0.2587 & -0.9364 \\ 0.4721 & 0 \end{bmatrix}, \\ B_5 &= \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}, B_6 = \begin{bmatrix} -1.6915 & 0 \\ 1.0249 & -0.3834 \end{bmatrix}. \end{aligned}$$

The transition probabilities between the scenarios are uniformly distributed, i.e., $\mathbb{P}(w = i) = 1/6$, $i \in \{1, 2, 3, 4, 5, 6\}$. Clearly there exists a switching sequence such that this open loop stochastic system is unstable. The objectives of the model predictive controller are to 1) guarantee closed-loop ULRSES, 2) satisfy the control input constraints, with $T_u = I_{2 \times 2}$, $u_{\max} = 2.5$, and 3) satisfy the state constraints, with $T_x = I_{2 \times 2}$, $x_{\max} = 5$. The initial state is $x_0(1) = x_0(2) = 2.5$. The objective cost function follows expression (9), with $Q = R = 0.01 \times I_{2 \times 2}$ and the one-step Markov polytopic risk metric is $\text{CVaR}_{0.75}$.

We simulated the state trajectories with 100 Monte Carlo samples, varying the number of lookahead steps N from 1 to 6, and compared the closed-loop performance from algorithms \mathcal{MPC} and \mathcal{MPC}^0 . Since at every time step we can only access the realizations of the stochastic system in the current simulation, we cannot compare the performance of the model predictive controller with Problem \mathcal{OPT} exactly. Instead, for each simulation, the MPC algorithm was run until a stage k' such that $\|x_{k'}\|_2 \leq 10^{-4}$. We then computed the *empirical* risk from all Monte Carlo simulations for a given horizon length using the cost function $J_{0,k'}$. The performance of the

MPC algorithms are evaluated based on the empirical risk and the cost upper bound (computed using Theorem VIII.1), and is summarized in Table I.

TABLE I: Performance for Example X-A.

Algorithms	Empirical Risk Cost (Upper Bound)	Mean (Variance) of Time per MPC Iteration
$C - \mathcal{MPC}^0$	4.016 (4.983)	Offline: 0.3574 (0.0091), Online: 0 (0)
$C - \mathcal{MPC}, N = 1$	2.882 (3.481)	Offline: 0.3252 (0.0023), Online: 0.1312 (0.0032)
$C - \mathcal{MPC}, N = 2$	1.686 (2.288)	Offline: 0.3241 (0.0042), Online: 0.8214 (0.0256)
$C - \mathcal{MPC}, N = 3$	1.105 (1.525)	Offline: 0.3380 (0.0133), Online: 2.9984 (0.3410)
$C - \mathcal{MPC}, N = 4$	0.898 (1.063)	Offline: 0.3421 (0.0117), Online: 40.9214 (2.4053)
$C - \mathcal{MPC}, N = 5$	0.676 (0.794)	Offline: 0.3989 (0.0091), Online: 498.9214 (15.4921)
$C - \mathcal{MPC}, N = 6$	0.440 (0.487)	Offline: 0.4011 (0.0154), Online: 7502.90075 (98.4104)

Solving the MPC problem with more lookahead steps decreases the performance index ($J_0^*(x_0)$), i.e., the sub-optimality gap of the MPC controller decreases. However, since the size of the online MPC problem scales exponentially with the number of lookahead steps, we can see that the online computation time scales exponentially from about 4 seconds at $N = 3$ to over 7300 seconds at $N = 6$. Due to this drastic increase in computation complexity, we are only able to run 5 Monte Carlo trials for each case at $N \in \{4, 5, 6\}$ for illustration. Note that the offline computation time is almost constant in all cases as the complexity of the offline problem is independent of the number of lookahead steps. Finally, using Theorem VIII.1 we obtain a lower bound value of 0.1276 for the optimal solution of problem \mathcal{OPT} . The looseness in the sub-optimality gap may be attributed to neglecting stability and state/control constraint guarantees in the lower bound derivation.

B. Comparison with [8]

Define the following stochastic system

$$A_1 = \begin{bmatrix} -0.8 & 1 \\ 0 & 0.8 \end{bmatrix}, A_2 = \begin{bmatrix} -0.8 & 1 \\ 0 & 1.2 \end{bmatrix}, A_3 = \begin{bmatrix} -0.8 & 1 \\ 0 & -0.4 \end{bmatrix},$$

$B_1 = B_2 = B_3 = [0, 1]^T$, the initial state is $x_0 = [5, 5]^T$, and uncertainty w_k is governed by an unknown probability mass function (different at each time step k), which belongs to the set of distributions

$$\mathcal{M} = \{m = \delta_1[0.5, 0.3, 0.2] + \delta_2[0.1, 0.6, 0.3] + \delta_3[0.2, 0.1, 0.7] : [\delta_1, \delta_2, \delta_3] \in \mathcal{B}\}.$$

TABLE II: Comparison with [8].

Method	Empirical Risk Cost	Standard Deviation	Mean (Variance) of Time per MPC Iteration
Algorithm \mathcal{MPC}	82.9913	78.0537	Offline: 0.1448 (0.0299), Online: 10.8312 (2.3914)
MPC algorithm in [8]	90.3348	98.0748	Offline: 0.1309 (0.0099), Online: 13.5374 (4.0046)

In this experiment we compare the performance of algorithm \mathcal{MPC} with a risk-averse adaptation of the MPC algorithm in [8] (the problem formulation is given in Appendix A). Notably, the cost function and stability constraints in [8] which only use the expectation operator, are replaced with their risk-sensitive counterparts, i.e., a time-consistent Markov polytopic risk measure (the one-step dynamic coherent risk in this example is defined as a distributionally robust expectation operator over the set distributions \mathcal{M} , i.e. $\rho(Z) = \max_{m \in \mathcal{M}} \mathbb{E}_m[Z]$). This was done to enable a fair comparison that is based solely on the *structure* of the overall solution algorithms and not on the objective or stability constraints.

The cost matrices used in this test are $Q = \text{diag}(1, 5)$ and $R = 1$. The state constraint matrix and threshold are given by $T_x = I_{2 \times 2}$, $x_{\max} = 12$, and the control constraint matrix and threshold are given by $T_u = 1$, $u_{\max} = 2$. While the MPC algorithm in [8] implemented scenario tree optimization techniques to reduce numerical complexity (to less than 20 nodes in their example), it is beyond the scope of this paper. For this reason, we choose $N = 3$ (giving 27 leaves in the scenario tree) to ensure that the above problems have similar online complexity. Table II shows the results from 100 Monte Carlo trials. Due to the additional complexity of the LMI conditions in algorithm \mathcal{MPC} , the offline computation time for our algorithm is slightly longer. Nevertheless, the resulting policy yields a lower empirical risk and standard deviation, and has a shorter online computation time as compared with its counterpart in [8]. This is due, in part, to the fact that we do not rely upon a fixed quadratic *robust* Lyapunov function (see Appendix A for further details) to guarantee closed-loop stability as in [8] and instead leverage the MPC cost function (see Theorem VII.4). This allows us to formulate a *less constrained* solution algorithm and achieve superior performance, accentuating the advantages of a risk averse approach to MPC.

C. Safety Brake in Adaptive Cruise Control

Adaptive cruise control (ACC) [56], [57] extends the functionalities of conventional cruise control. In addition to tracking the reference velocity of the driver, ACC also enforces a separation

distance between the leading vehicle (the host) and the follower (the vehicle that is equipped with the ACC system) to improve passenger comfort and safety. This crucial safety feature prevents a car crash when the host stops abruptly due to unforeseeable hazards.

In this experiment, we design a risk-sensitive controller for the ACC system that guarantees a safe separation distance between vehicles even when the host stops abruptly. As a prediction model for the MPC control problem, we define v_k and a_k to be the speed and the acceleration of the follower respectively, and $v_{l,k}$, $a_{l,k}$ as the velocity and acceleration of the leader. The acceleration a_k is modeled as the integrator

$$a_{k+1} = a_k + T_s u_k,$$

where T_s is the sampling period, and the control input u_k is the rate of change of acceleration (jerk) which is assumed to be constant over the sampling interval. The leader and follower velocities are given by

$$v_{k+1} = v_k + T_s a_k, \quad v_{l,k+1} = w_k v_{l,k},$$

where w_k is the leader's geometric deceleration rate. Since this rate captures the degree of abrupt stopping, its evolution has a stochastic nature. Here we assume w_k belongs to the sample space $\mathcal{W} = \{0.5, 0.7, 0.9\}$ whose transition follows a uniform distribution. Furthermore, the distance d_k between the leader and the follower evolves as

$$d_{k+1} = d_k + T_s (v_{l,k} - v_k).$$

In order to ensure safety, we also set the reference distance to be velocity dependent, which can be modeled as $d_{\text{ref},k} = \delta_{\text{ref}} + \gamma_{\text{ref}} v_k$ with $\delta_{\text{ref}} = 4m$ and $\gamma_{\text{ref}} = 3s$. Together, the system dynamics may be written as (2), where $w_k \in \{1, 2, 3\}$, and $x_k := [d_k - d_{\text{ref},k}, v_k, a_k, v_{l,k}]$.

In order to guarantee comfort and safety, we assume the constraints $|u_k| \leq 3ms^{-3}$ (bounded jerk), and $|v_k| \leq 12ms^{-1}$ (bounded speed), and the state and control weighting matrices within the quadratic cost are $Q = \text{diag}(Q_d, Q_v, 0, 0)$, $R = Q_u$, where Q_d , Q_v , Q_u are the weights on the separation distance tracking error, velocity, and jerk, respectively. To study the risk-averse behavior of the safety brake mechanism, we design a risk-sensitive MPC controller based on the dynamic risk compounded by the mean absolute semi-deviation with $c = 1$. For demonstrative purposes, the MPC lookahead step is simply set to one ($N = 1$). The performance of the risk sensitive ACC system is illustrated by the state trajectories in Figure 2 (for brevity we only show the distance trajectory d_k). It can be seen that the controller is able to stabilize

TABLE III: Statistics for Risk-Sensitive ACC System.

Method	Mean Cost	Standard Deviation	Mean (Variance) of Time per MPC Iteration
$c = 1$	451.3442	9.5854	Offline: 0.3944 (0.0036) , Online: 0.0727 (0.0054)
$c = 0$ (Risk Neutral)	423.8701	12.7447	Offline: 0.4011 (0.0024) , Online: 0.0450 (0.0067)

the stochastic error $d_k - d_{\text{ref},k}$ (in the risk-sensitive sense) such that the speed of the follower vehicle gradually vanishes, and the separation distance d_k between the two cars converges to the constant δ_{ref} . Notice that besides error tracking, the dynamic mean semi-deviation risk sensitive objective function also regulates the variability of distance separation between the two vehicles, as shown in Figure 2 and Table III. Compared with the risk-neutral MPC approach ($c = 0$), this risk-sensitive ACC system results in a lower variance in separation distance, suggesting a more comfortable passenger experience.

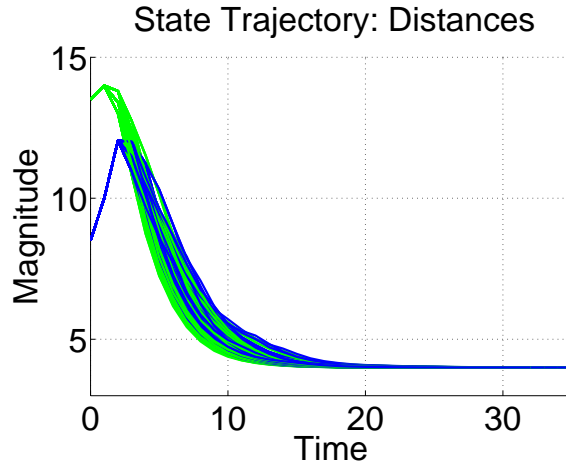


Fig. 2: Separation distance d_k trajectories of the Safety Brake in Risk Sensitive Adaptive Cruise Control.

XI. CONCLUSION AND FUTURE WORK

In this paper we presented a framework for risk-averse MPC by leveraging recent advances in the theory of dynamic risk metrics developed by the operations research community. The proposed approach has the following advantages: (1) it is axiomatically justified and leads to time-consistent risk assessments; (2) it is amenable to dynamic and convex programming; and (3) it is general, in that it captures a full range of risk assessments from risk-neutral to worst

case (due to the generality of Markov polytopic risk metrics). Our framework thus provides a unifying perspective on risk-sensitive MPC.

We plan to extend our work to handle cases where the state and control constraints are required to hold only with a given probability threshold (in contrast to hard constraints) by exploiting techniques such as *probabilistic invariance* [58]. This relaxation has the potential to provide significantly improved performance at the risk of occasionally violating constraints. Second, we plan to combine our approach with methods for scenario tree optimization in order to reduce the online computation load. Third, while polytopic risk metrics encompass a wide range of possible risk assessments, extending our work to non-polytopic risks and more general stage-wise costs can broaden the domain of application of the approach. Finally, an important consideration from a practical standpoint is the *choice* of risk metric appropriate for a given application. We plan to develop principled approaches for making this choice, e.g., by computing polytopic risk envelopes based on confidence regions for the disturbance model.

APPENDIX A

ALTERNATIVE FORMULATION OF PROBLEM \mathcal{PE} AND \mathcal{MPC}

In this section we present alternative formulations of problems \mathcal{PE} and \mathcal{MPC} inspired by the approach in [8]. The methodology here is to design (offline) an equivalent control invariant set \mathcal{E}_{\max} and a *robust* Lyapunov function such that ULRSES and constraint fulfillment are guaranteed using a local state feedback control law $u(x) = Fx$. Let $P = P^\top \succ 0$ and $L = L^\top \succ 0$. Define $V(x_k) = x_k^\top P x_k$. If

$$V(x_{k+1}) - V(x_k) \leq -x_k^\top L x_k, \quad \text{surely, } \forall k \in \mathbb{N}, \quad (36)$$

then $V(x_k)$ is a *robust* Lyapunov function for system (2). In the online problem, the inequality above is relaxed to its stochastic counterpart as shown in (8). We first formalize the offline optimization problem:

Optimization Problem \mathcal{PE} — Given an initial state $x_0 \in \mathbb{X}$, and a matrix $L = L^\top \succ 0$, solve

$$\begin{aligned} & \max_{W=W^\top \succ 0, Y, \gamma > 0} \quad \text{logdet}(W) \\ & \text{subject to} \quad x_0^\top W^{-1} x_0 \leq 1, \\ & \quad Y^\top \frac{T_u^\top T_u}{u_{\max}^2} Y \preceq W, \\ & \quad (A_j W + B_j Y)^\top \frac{T_x^\top T_x}{x_{\max}^2} (A_j W + B_j Y) \preceq W, \forall j \in \{1, \dots, L\}, \\ & \quad \begin{bmatrix} W & (L^{1/2} W)^\top & (A_j W + B_j Y)^\top \\ * & \gamma I_{N_x} & 0 \\ * & 0 & W \end{bmatrix} \succeq 0, \forall j \in \{1, \dots, L\}. \end{aligned} \quad (37)$$

Suppose problem \mathcal{PE} above is feasible. Set $P = \gamma W^{-1}$. The control invariant set is then defined to be the intersection $\mathbb{X} \cap \mathcal{E}_{\max}$, where $\mathcal{E}_{\max} := \{x \in \mathbb{R}^{N_x} \mid x^\top W^{-1} x \leq 1\} = \{x \in \mathbb{R}^{N_x} \mid x^\top P x \leq \gamma\}$.

Note that $\mathbb{X} \cap \mathcal{E}_{\max}$ is a robust control invariant set under the feasible local feedback control law $u(x) = YW^{-1}x$. The constraint in (37) is an equivalent reformulation of the robust Lyapunov condition given in (36) where $x_{k+1} = (A(w) + B(w)YG^{-1})x$. That is, the closed-loop dynamics are ULRSES with domain $\mathbb{X} \cap \mathcal{E}_{\max}$ under the feedback control law $u(x) = YG^{-1}x$. In an attempt to improve the stability properties of the system beyond what is achievable via this feedback control law, the online MPC problem is formalized as follows:

Optimization problem \mathcal{MPC} — Given an initial state $x_{k|k} \in \mathbb{X} \cap \mathcal{E}_{\max}$ and a prediction horizon $N \geq 1$, solve

$$\begin{aligned} \min_{\pi_{k|k}, \dots, \pi_{k+N-1|k}} \quad & J(x_{k|k}, \pi_{k|k}, \dots, \pi_{k+N-1|k}, P) \\ \text{subject to} \quad & x_{k+h+1|k} = A(w_{k+h})x_{k+h|k} + B(w_{k+h})\pi_{k+h|k}(x_{k+h|k}), \\ & \pi_{k+h|k}(x_{k+h|k}) \in \mathbb{U}, x_{k+h+1|k} \in \mathbb{X}, h \in \{0, \dots, N-1\}, \\ & x_{k+1|k} \in \mathcal{E}_{\max}(W) \text{ surely,} \end{aligned} \tag{38}$$

$$\rho_k \left((Ax_{k|k} + B\pi_{k|k})^\top P (Ax_{k|k} + B\pi_{k|k}) \right) - x_k^\top P x_k \leq -x_{k|k}^\top L x_{k|k}. \tag{39}$$

Provided problem \mathcal{MPC} is recursively feasible, ULRSSES with domain $\mathbb{X} \cap \mathcal{E}_{\max}$ is enforced automatically via (39) which leverages the risk-sensitive Lyapunov function $x_k^\top P x_k$, where P is the solution to the offline problem. Persistent feasibility however, is guaranteed by the constraint in (38).

A few remarks are in order. First, notice that problem \mathcal{PE} also includes an initial condition inclusion constraint $x_0 \in \mathcal{E}_{\max}$. That is, the set \mathcal{E}_{\max} is *re-designed* for each initial condition in order to ensure the *existence* of a feasible risk-sensitively stabilizing control law, and thereby the feasibility of problem \mathcal{MPC} . Second, constraint (38), necessary for ensuring recursive feasibility for problem \mathcal{MPC} , is rather restrictive as compared with how the control invariant set \mathcal{E}_{\max} is used within our algorithm, i.e., as a terminal set constraint. When combined with the risk-sensitive stability constraint (39), necessary due to the reliance upon a *fixed* quadratic Lyapunov function, the resulting formulation yields worse performance (in terms of cost function optimality), and a smaller domain of feasibility.

APPENDIX B

PROOF OF THEOREM IX.1 AND COROLLARY IX.2

We first present the Projection Lemma:

Lemma B.1 (Projection Lemma). *For matrices $\Omega(X)$, $U(X)$, $V(X)$ of appropriate dimensions, where X is a matrix variable, the following statements are equivalent:*

1) *There exists a matrix W such that*

$$\Omega(X) + U(X)WV(X) + V(X)^\top W^\top U(X)^\top \prec 0.$$

2) *The following inequalities hold:*

$$U(X)^\perp \Omega(X) (U(X)^\perp)^\top \prec 0, \quad (V(X)^\top)^\perp \Omega(X) ((V(X)^\top)^\perp)^\top \prec 0,$$

where A^\perp is the orthogonal complement of A .

Proof. See Chapter 2 in [50]. □

We now give the proof for Theorem IX.1 by leveraging the Projection lemma:

Proof. (Proof of Theorem IX.1) Using simple algebraic factorizations, $\forall l \in \{1, \dots, \text{cardinality}(\mathcal{U}^{\text{poly}, V}(p))\}$, inequality (20) can be rewritten as

$$\begin{bmatrix} I \\ \Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ F \\ Q^{\frac{1}{2}} \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 & 0 \\ 0 & -I_{L \times L} \otimes P & 0 & 0 \\ 0 & 0 & -R & 0 \\ 0 & 0 & 0 & -I \end{bmatrix} \begin{bmatrix} I \\ \Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ F \\ Q^{\frac{1}{2}} \end{bmatrix} \succ 0.$$

By Schur complement, the above expression is equivalent to

$$\begin{bmatrix} I & 0 & 0 & \Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ 0 & I & 0 & F \\ 0 & 0 & I & Q^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} I_{L \times L} \otimes \bar{Q} & 0 & 0 & 0 \\ 0 & R^{-1} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -\bar{Q} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ (\bar{A} + \bar{B}F)^\top \Sigma_l^{\frac{1}{2}} & F^\top & Q^{\frac{1}{2}} \end{bmatrix} \succ 0,$$

where $\bar{Q} = P^{-1}$. Now since $\bar{Q} = \bar{Q}^\top \succ 0$ and $R = R^\top \succ 0$, we also have the following identity:

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} I_{L \times L} \otimes \bar{Q} & 0 & 0 & 0 \\ 0 & R^{-1} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -\bar{Q} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix} \succ 0.$$

Next, notice that

$$\begin{bmatrix} -\Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ -F \\ -Q^{\frac{1}{2}} \\ I \end{bmatrix}^\perp = \begin{bmatrix} I & 0 & 0 & \Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ 0 & I & 0 & F \\ 0 & 0 & I & Q^{\frac{1}{2}} \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ I \end{bmatrix}^\perp = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix}.$$

Now, set:

$$\Omega = - \begin{bmatrix} I_{L \times L} \otimes \bar{Q} & 0 & 0 & 0 \\ 0 & R^{-1} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -\bar{Q} \end{bmatrix}, \quad U = \begin{bmatrix} -\Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ -F \\ -Q^{\frac{1}{2}} \\ I \end{bmatrix}, \quad V^T = \begin{bmatrix} 0 \\ 0 \\ 0 \\ I \end{bmatrix}.$$

Then by Lemma B.1, it is equivalent to find a matrix G that satisfies the following inequality

$\forall l \in \{1, \dots, \text{cardinality}(\mathcal{U}^{\text{poly}, V}(p))\}$:

$$\begin{bmatrix} I_{L \times L} \otimes \bar{Q} & 0 & 0 & 0 \\ 0 & R^{-1} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -\bar{Q} \end{bmatrix} + \begin{bmatrix} -\Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ -F \\ -Q^{\frac{1}{2}} \\ I \end{bmatrix} G \begin{bmatrix} 0 \\ 0 \\ 0 \\ I \end{bmatrix}^\top + \begin{bmatrix} 0 \\ 0 \\ 0 \\ I \end{bmatrix} G^\top \begin{bmatrix} -\Sigma_l^{\frac{1}{2}}(\bar{A} + \bar{B}F) \\ -F \\ -Q^{\frac{1}{2}} \\ I \end{bmatrix}^\top \succ 0. \quad (40)$$

Setting $F = YG^{-1}$ and pre-and post-multiplying the above inequality by $\text{diag}(I, R^{\frac{1}{2}}, I, I)$ yields the LMI given in (29). Furthermore, from the inequality $-\bar{Q} + G + G^\top \succ 0$ where $\bar{Q} \succ 0$, we know that $G + G^\top \succ 0$. Thus, by the Lyapunov stability theorem, the linear time-invariant system $\dot{x} = -Gx$ with Lyapunov function $x^\top x$ is asymptotically stable (i.e. all eigenvalues of G have positive real part). Therefore, G is an invertible matrix and $F = YG^{-1}$ is well defined. \square

Proof. (Proof of Corollary IX.2) We will prove that the third inequality in (30) implies inequality (22). Details of the proofs on the implications of the first two inequalities in (30) follow from identical arguments and will be omitted for the sake of brevity. Using simple algebraic factorizations, inequality (22) may be rewritten (in strict form) as:

$$\begin{bmatrix} I \\ A_j + B_j F \end{bmatrix}^\top \begin{bmatrix} W^{-1} & 0 \\ 0 & -W^{-1} \end{bmatrix} \begin{bmatrix} I \\ A_j + B_j F \end{bmatrix} \succ 0, \quad \forall j \in \{1, \dots, L\}.$$

By Schur complement, the above expression is equivalent to

$$\begin{bmatrix} I & A_j + B_j F \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & -W \end{bmatrix} \begin{bmatrix} I \\ (A_j + B_j F)^\top \end{bmatrix} \succ 0, \quad \forall j \in \{1, \dots, L\}.$$

Furthermore since $W \succ 0$, we also have the identity

$$\begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & -W \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} \succ 0.$$

Now, notice that:

$$\begin{bmatrix} -(A_j + B_j F) \\ I \end{bmatrix}^\perp = \begin{bmatrix} I & A_j + B_j F \end{bmatrix}, \quad \begin{bmatrix} 0 \\ I \end{bmatrix}^\perp = \begin{bmatrix} I & 0 \end{bmatrix}.$$

Then by Lemma B.1, it is equivalent to find a matrix G such that the following inequality holds

$\forall j \in \{1, \dots, L\}$:

$$\begin{bmatrix} W & 0 \\ 0 & -W \end{bmatrix} + \begin{bmatrix} -(A_j + B_j F) \\ I \end{bmatrix} G \begin{bmatrix} 0 \\ I \end{bmatrix}^\top + \begin{bmatrix} 0 \\ I \end{bmatrix} G^\top \begin{bmatrix} -(A_j + B_j F) \\ I \end{bmatrix}^\top \succ 0. \quad (41)$$

Note that Lemma B.1 provides an equivalence (necessary and sufficient) condition between (41) and (22) if G is allowed to be any arbitrary LMI variable. However, in order to restrict G to be the same variable as in Theorem IX.1, the equivalence relation reduces to sufficiency only. Setting $F = YG^{-1}$ in the above expression gives the claim. \square

REFERENCES

- [1] J. Qin and T. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.
- [2] Y. Wang and S. Boyd, "Fast model predictive control using online optimization," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 2, pp. 267–278, 2010.
- [3] M. V. Kothare, V. Balakrishnan, and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, vol. 32, no. 10, pp. 1361–1379, 1996.
- [4] C. E. de Souza, "Robust stability and stabilization of uncertain discrete-time Markovian jump linear systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 5, pp. 836–841, 2006.
- [5] J. Löfberg, "Minimax approaches to robust model predictive control," Ph.D. dissertation, Linköping University, 2003.
- [6] D. Q. Mayne, "Model predictive control: Recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.
- [7] J. A. Primbs and C. H. Sung, "Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 221–230, 2009.
- [8] D. Bernardini and A. Bemporad, "Stabilizing model predictive control of stochastic constrained linear systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1468–1480, 2012.
- [9] M. Cannon, B. Kouvaritakis, S. V. Rakovic, and Q. Cheng, "Stochastic tubes in model predictive control with probabilistic constraints," *IEEE Transactions on Automatic Control*, vol. 56, no. 1, pp. 194–200, 2011.
- [10] J. Fleming, M. Cannon, and B. Kouvaritakis, "Stochastic tube MPC for LPV systems with probabilistic set inclusion conditions," in *Proc. IEEE Conf. on Decision and Control*, 2014.

- [11] D. Mayne, “Robust and stochastic model predictive control: Are we going in the right direction?” *Annual Reviews in Control*, vol. 41, pp. 184–192, 2016.
- [12] A. Mesbah, “Stochastic model predictive control: An overview and perspectives for future research,” *IEEE Control Systems Magazine*, vol. 36, no. 6, pp. 30–44, 2016.
- [13] S. Mannor and J. N. Tsitsiklis, “Mean-variance optimization in Markov decision processes,” in *Int. Conf. on Machine Learning*, 2011.
- [14] A. Sarlette, “Geometry and symmetries in coordination control,” Ph.D. dissertation, Université of Liège, Belgium, 2009.
- [15] B. Roorda, J. M. Schumacher, and J. Engwerda, “Coherent acceptability measures in multi-period models,” *Mathematical Finance*, vol. 15, no. 4, pp. 589–612, 2005.
- [16] A. Ruszczyński, “Risk-averse dynamic programming for Markov decision process,” *Mathematical Programming*, vol. 125, no. 2, pp. 235–261, 2010.
- [17] H. Markowitz, “Portfolio selection,” *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [18] H. U. Gerber and G. Pafum, “Utility functions: From risk theory to finance,” *North American Actuarial Journal*, vol. 2, no. 3, pp. 74–91, 1998.
- [19] W. H. Fleming and S. J. Sheu, “Risk-sensitive control and an optimal investment model,” *Mathematical Finance*, vol. 10, no. 2, pp. 197–213, 2000.
- [20] R. Howard and J. Matheson, “Risk-sensitive Markov decision processes,” *Management Science*, vol. 8, no. 7, pp. 356–369, 1972.
- [21] O. Mihatsch and R. Neuneier, “Risk-sensitive reinforcement learning,” *Machine Learning*, vol. 49, no. 2, pp. 267–290, 2002.
- [22] N. Bäuerle and U. Rieder, “More risk-sensitive Markov decision processes,” *Mathematics of Operations Research*, vol. 39, no. 1, pp. 105–120, 2013.
- [23] P. Whittle, “Risk-sensitive linear/quadratic/gaussian control,” *Advances in Applied Probability*, vol. 13, no. 4, pp. 764–777, 1981.
- [24] K. Glover and J. C. Doyle, “Relations between H_∞ and risk sensitive controllers,” in *Analysis and Optimization of Systems*. Springer-Verlag, 1987.
- [25] K. Detlefsen and G. Scandolo, “Conditional and dynamic convex risk measures,” *Finance and Stochastics*, vol. 9, no. 4, pp. 539–561, 2005.
- [26] R. C. Green and B. Hollifield, “When will mean-variance efficient portfolios be well diversified?” *Journal of Finance*, vol. 47, no. 5, pp. 1785–1809, 1992.
- [27] M. J. Best and R. R. Grauer, “On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results,” *The Review of Financial Studies*, vol. 4, no. 2, pp. 315–342, 1991.
- [28] H. Föllmer and T. Knispel, “Entropic risk measures: Coherence vs convexity, model ambiguity and robust large deviations,” *Stochastics and Dynamics*, vol. 11, no. 02n03, pp. 333–351, 2011.
- [29] M. Rabin, “Risk aversion and expected-utility theory: A calibration theorem,” *Econometrica*, vol. 68, no. 5, pp. 1281–1292, 2000.
- [30] C. X. Wang, S. Webster, and N. C. Suresh, “Would a risk-averse newsvendor order less at a higher selling price?” *European Journal of Operational Research*, vol. 196, no. 2, pp. 544–553, 2009.
- [31] A. Ruszczyński and A. Shapiro, “Optimization of convex risk functions,” *Mathematics of Operations Research*, vol. 31, no. 3, pp. 433–452, 2006.
- [32] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, “Coherent measures of risk,” *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

- [33] Y. Chow and M. Pavone, "A framework for time-consistent, risk-averse model predictive control: Theory and algorithms," in *American Control Conference*, 2014.
- [34] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming: Modeling and theory*, 2nd ed. SIAM, 2014.
- [35] P. Cheridito and M. Stadjé, "Time-inconsistency of VaR and time-consistent alternatives," *Finance Research Letters*, vol. 6, no. 1, pp. 40–46, 2009.
- [36] A. Shapiro, "Dynamic programming approach to adjustable robust optimization," *Operations Research Letters*, vol. 39, no. 2, pp. 83–87, 2010.
- [37] D. A. Iancu, M. Petrik, and D. Subramanian, "Tight approximations of dynamic risk measures," *Mathematics of Operations Research*, vol. 40, no. 3, pp. 655–682, 2015.
- [38] A. Eichhorn and W. Römisch, "Polyhedral risk measures in stochastic programming," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 69–95, 2005.
- [39] W. Ogryczak and A. Ruszczyński, "From stochastic dominance to mean-risk models: Semi-deviations as risk measures," *European Journal of Operational Research*, vol. 116, no. 1, pp. 33–50, 1999.
- [40] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, pp. 21–41, 2000.
- [41] D. Bertsimas and D. B. Brown, "Constructing uncertainty sets for robust linear optimization," *Operations Research*, vol. 57, no. 6, pp. 1483–1495, 2009.
- [42] A. Ben-Tal and M. Teboulle, "An old-new concept of convex risk measure: The optimized certainty equivalent," *Mathematical Finance*, vol. 17, no. 3, pp. 449–476, 2007.
- [43] L. B. Ryashko and H. Schurz, "Mean square stability analysis of some linear stochastic systems," *Dynamic Systems and Applications*, vol. 6, no. 2, pp. 165–190, 1996.
- [44] C. W. Scherer, "Mixed H_2/H_∞ control for time-varying and linear parametrically-varying systems," *Int. Journal of Robust and Nonlinear Control*, vol. 6, no. 9-10, pp. 929–952, 1996.
- [45] M. Chilali and P. Gahinet, " H_∞ design with pole placement constraints: an LMI approach," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 358–367, 1996.
- [46] V. A. Yakubovich, "The S-procedure in non-linear control theory," *Vestnik Leningrad University Math*, vol. 4, pp. 73–93, 1977, In Russian, 1971.
- [47] J. Rawlings and D. Mayne, *Model predictive control: Theory and design*. Nob Hill Publishing, 2013.
- [48] B.-G. Park and W. H. Kwon, "Robust one-step receding horizon control of discrete-time markovian jump uncertain systems," *Automatica*, vol. 38, no. 7, pp. 1229–1235, 2002.
- [49] O. Toker and H. Ozbay, "On the NP-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback," in *American Control Conference*, 1995.
- [50] R. E. Skelton, T. Iwasaki, and K. Grigoriadis, *A Unified Algebraic Approach to Linear Control Design*. CRC Press, 1997.
- [51] C. W. Scherer, "The general nonstrict algebraic Riccati inequality," *Linear Algebra and its Applications*, vol. 19, pp. 1–33, 1995.
- [52] Y. Chow and M. Pavone, "A uniform-grid discretization algorithm for stochastic optimal control with risk constraints," in *Proc. IEEE Conf. on Decision and Control*, 2013.
- [53] C.-S. Chow and J. N. Tsitsiklis, "An optimal one-way multigrid algorithm for discrete-time stochastic control," *IEEE Transactions on Automatic Control*, vol. 36, no. 8, pp. 898–914, 1991.
- [54] M. Herceg, M. Kvasnica, C. N. Jones, and M. Morari, "Multi-parametric toolbox 3.0," in *European Control Conference*, 2013.

- [55] J. Löfberg, “YALMIP : A toolbox for modeling and optimization in MATLAB,” in *IEEE Int. Symp. on Computer Aided Control Systems Design*, 2004.
- [56] C.-Y. Liang and H. Peng, “Optimal adaptive cruise control with guaranteed string stability,” *Vehicle System Dynamics*, vol. 32, no. 4-5, pp. 313–330, 1999.
- [57] M. Bichi, G. Ripaccioli, S. Di Cairano, D. Bernardini, A. Bemporad, and I. V. Kolmanovsky, “Stochastic model predictive control with driver behavior learning for improved powertrain control,” in *Proc. IEEE Conf. on Decision and Control*, 2010.
- [58] M. Cannon, B. Kouvaritakis, and X. Wu, “Probabilistic constrained mpc for multiplicative and additive stochastic uncertainty,” *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1626–1632, 2009.



Yin-Lam Chow received a B. Eng degree in Mechanical Engineering from the University of Hong Kong in 2009, a M.Sc. degree in Aerospace Engineering from Purdue University in 2011, and is currently a Ph.D. candidate at Stanford University in the Institute for Computational and Mathematical Engineering. His research interests include control theory, machine learning, sequential decision making and reinforcement learning.



Sumeet Singh is a Ph.D. candidate at Stanford University in the Autonomous Systems Lab. He received a B. Eng. in Mechanical Engineering and a Diploma of Music (Performance) from the University of Melbourne in 2012, and a M.Sc. in Aeronautics and Astronautics from Stanford University in 2015. Prior to joining Stanford, Sumeet worked in the Berkeley Micromechanical Analysis and Design lab at University of California Berkeley in 2011, and the Aeromechanics Branch at NASA Ames in 2013. Sumeet's current research interests include the design of stochastic and robust nonlinear MPC algorithms, with specific applications to spacecraft control.



Anirudha Majumdar is currently a postdoctoral scholar at Stanford University in the Autonomous Systems Lab. He will be starting as an Assistant Professor at Princeton University in 2017. He completed his Ph.D. in the Electrical Engineering and Computer Science department at MIT. Majumdar received his undergraduate degree in Mechanical Engineering and Mathematics at the University of Pennsylvania, where he was a member of the GRASP lab. His research interests lie in controlling highly agile robotic systems such as UAVs with formal safety guarantees. Majumdar's research has been recognized with the Siebel Foundation Scholarship and the Best Conference Paper Award at the International Conference on Robotics and Automation (ICRA) 2013.



Marco Pavone is an Assistant Professor of Aeronautics and Astronautics at Stanford University, where he is the Director of the Autonomous Systems Laboratory. Before joining Stanford, he was a Research Technologist within the Robotics Section at the NASA Jet Propulsion Laboratory. He received a Ph.D. degree in Aeronautics and Astronautics from the Massachusetts Institute of Technology in 2010. Dr. Pavone's areas of expertise lie in the fields of controls and robotics. His main research interests are in the development of methodologies for the analysis, design, and control of autonomous systems, with an emphasis on autonomous aerospace vehicles and large-scale robotic networks. He is a recipient of a PECASE Award, an ONR YIP Award, an NSF CAREER Award, a NASA Early Career Faculty Award, a Hellman Faculty Scholar Award, and was named NASA NIAC Fellow in 2011. He is currently serving as an Associate Editor for the IEEE Control Systems Magazine.